Westat®

Improving lives through research™

Review of the FoodAPS 2012 Instrument Design, Response Burden, Use of Incentives, and Response Rates

Authors

Ting Yan
Aaron Maitland

December 12, 2016

Prepared for:
Economic Research Service
U.S. Department of Agriculture
355 E Street, SW
Washington, DC 20024-3221

Prepared by:
Westat
An Employee-Owned Research Corporation®
1600 Research Boulevard
Rockville, Maryland 20850-3129
(301) 251-1500

Preferred citation:

Yan, T., and Maitland, A. (2016). *Review of the FoodAPS 2012 Instrument Design, Response Burden, Use of Incentives, and Response Rates.* Prepared for the Economic Research Service, U.S. Department of Agriculture. Washington, D.C.

This report is part of a series of five reports. The citations for the other reports are as follows:

Krenzke, T., and Kali, J. (2016). *Review of the FoodAPS 2012 Sample Design.* Prepared for the Economic Research Service, U.S. Department of Agriculture. Washington, D.C.

Li, J., Van de Kerckhove, W., and Krenzke, T. (2016). *Review of the FoodAPS 2012 Imputation Approaches for Income and Price Data.* Prepared for the Economic Research Service, U.S. Department of Agriculture. Washington, D.C.

Maitland, A., and Li, L. (2016). *Review of the Completeness and Accuracy of FoodAPS 2012 Data.* Prepared for the Economic Research Service, U.S. Department of Agriculture. Washington, D.C.

Petraglia, E., Van de Kerckhove, W., and Krenzke, T. (2016). *Review of the Potential for Nonresponse Bias in FoodAPS 2012.* Prepared for the Economic Research Service, U.S. Department of Agriculture. Washington, D.C.

Table of Contents

Executive Summary

The 2012 National Household Food Acquisition and Purchase Survey (FoodAPS) (hereafter referred to as "FoodAPS-1") is a household survey fielded primarily in 2012 and designed to capture detailed information on the food acquisitions of U.S. households. FoodAPS-1 was sponsored by the U.S. Department of Agriculture and managed by its Economic Research Service (ERS). In 2015, ERS contracted with Westat to conduct an independent assessment of the quality of the FoodAPS-1 sample design, instrumentation, data collection procedures, and resulting data. This report is part of a series of five reports that constitute that assessment.

FoodAPS-1 collects comprehensive data on American households' food acquisition, factors influencing food choices, and household well-being. The data collection includes a household screening form; two in-person interview visits (one at the beginning of the reference week and one at the end of the reference week); a paper Income Worksheet, three paper food forms (Primary Respondent Food Form, Adult/Youth Food Form, Meals and Snacks Form); three telephone interviews intended to be initiated by sample households to report their food acquisitions; and a paper Feedback Form to be completed by sample households throughout the study week. The overall weighted response rate was 41.5 percent.

Westat assessed all instruments and study materials employed in FoodAPS-1 against best practices in the survey field and the latest empirical evidence in the survey literature. This report describes our review of the instruments and forms, and presents our recommendations for each of the instruments and forms. We summarize the key findings here.

Household Screening Form. FoodAPS-1 employed a paper-pencil form for the screening interview to determine household eligibility. Various issues with this paper-pencil form have been documented, such as missing data for key variables (e.g., personal ID, Q8 used to determine household size), duplicate forms with the same ID, and missing forms; these issues can be largely prevented or mitigated by computerizing the form. Our expert review of the paper-pencil screener form indicates that two sets of questions need further revision to better suit ERS' analytic goals – questions on household size (Q7 to Q8a) and questions on household income (Q9 to Q11). Westat suggests new wordings for those two sets of questions.

Initial In-Person Interview. FoodAPS-1 conducted the Initial Interview of sampled households via Computer-Assisted Personal Interviewing (CAPI). Westat found that the full potential of a

computerized instrument was not taken advantage of, leading to issues such as missing data, outlying values, and inconsistent answers. Westat recommends adding edits to the CAPI instrument to reduce the above-mentioned data issues. In addition, questions on household roster (A1 to A3c3) do not meet ERS' analytic goals. These questions are not clear as to whom to include and whom to exclude and, consequently, result in an undercoverage of people who are not living in the household but contribute substantially to the household income or incur substantial expenses. The Supplemental Nutrition Assistance Program (SNAP) participation questions (Q1 to B3a) are problematic. Q1 is intended to capture whether or not households currently receive SNAP. For those who responded "Yes," they are followed up with B1, which asks about the date of their last receipt. For those who answered "No," "Don't Know," or "Refused" to Q1, they are followed up with B3, which asks whether or not they ever received SNAP. B3a is intended to capture the date of last receipt. However, the wording and the order of these questions resulted in some respondents saying "Yes" to the current receipt question (Q1) but providing a date older than 30 days ago (at B1) and others saying "No" to Q1 and then providing a date within the last 30 days. Additionally, these questions do not allow ERS to identify SNAP units within a household. Westat proposed alternative question wordings and alternative sequence to address these issues.

Final In-Person Interview. Similar to the Initial Interview, the Final Interview is also conducted via CAPI, and the full potential of the computerization of the instrument was not utilized. In terms of specific questions asked in the Final Interview, we find a few issues. First, the wordings of the income questions (F1 to F8b) are prone to misunderstanding. Second, the F4 series fails to identify Temporary Assistance for Needy Families (TANF) as a separate source of income. Third, the current instrument does not capture the SNAP benefits that the household may receive during the data collection week. Westat recommends altering the question to capture information regarding the week the household receives SNAP benefits.

Income Worksheet. The Income Worksheet is an optional paper form provided to the primary respondents to be completed during the data collection week in preparation for the Final Interview. The low response rate to the worksheet at FoodAPS-1 diminishes the utility of the worksheet. Westat again recommends computerizing the form and to provide more training at the end of the Initial Interview to educate respondents and encourage them to fill out the form.

Food Acquisition Forms. FoodAPS-1 asked all household members age 11 years and older to track and report all food acquisitions through food books during the 1-week period. Three different types of food books were used. The primary respondent food book was intended for primary respondents, who were responsible for recording all food-at-home (FAH) items acquired by any

household members, recording their own acquisition of food-away-from-home (FAFH) items, and recording FAFH items acquired by members less than 11 years old. The primary respondent food book includes Daily List pages, Red Pages (to record FAFH items), Blue Pages (to record FAH items), and Barcode Pages. The adult food book is intended for adult household members aged 19 or above, and it only contains Daily List pages and Red Pages. The adult household members were instructed to record FAFH items in their own food book (on Red Pages) and enter FAH items into the Blue Pages of the primary food book. The youth food book was for household members age 11 or above and contained only Red Pages. Youths were instructed to record their acquisition of FAFH items on their own food books.

Food acquisitions reported in the books were key-entered either by data entry coders or telephone interviewers and required extensive manual review, which took time and resources. Westat recommends that ERS consider an alternative reporting system with the potential to remove the need for manual entry of data, allow automatic near-real-time checking and review of the data, and enable timely feedback to respondents when, for instance, inconsistent entries are identified. An alternative reporting system should also have the potential to make it easier for respondents to enter item descriptions for both FAH and FAFH purchases.

FoodAPS-1 asked the primary respondents to use a scanner to scan barcodes on FAH food or drink items. Westat evaluated the performance of the scanner and found that over 75 percent of the items with scanned barcodes can be matched to the Universal Product Code (UPC) data dictionary for product names. However, after assigning a product name based on receipts, there are still about 11.2 percent of total food items reported that do not have item descriptions. Westat recommends providing timely feedback on whether the scanned barcodes matches to product name with an item description so that if the item description is missing, respondents can be instructed to manually enter it.

The Meals and Snacks Form is a paper form to be completed by the primary respondents during the data collection week. Westat recommends computerizing and pre-filling this form with the household member roster and days of the week to reduce the extent of editing and key-entry that was experienced in FoodAPS-1.

Telephone Interview for Reporting Food Acquisitions. Westat also reviewed the instrument used for telephone interviews. If ERS decides to continue with the Telephone Interviews as in FoodAPS-1, Westat suggests that ERS ensure that the programing includes logic edits to check for potential redundancy (e.g., a meal shared by all household members at a restaurant is reported on

multiple food books); inconsistency (e.g., the same meal was reported with different dollar amounts at the Daily Lists section and the Red Pages); and out-of-range values (e.g., a meal purchased at McDonald's costs more than $100) in reports of food acquisitions. Westat also recommends that ERS consider using an alternative reporting system to replace the telephone interviews.

Feedback Form. The Respondent Feedback Form is a paper form that is given to the respondent during the Final Visit. Respondents are asked to complete the form and enclose it in a sealed envelope. Westat suggests that ERS administer the Respondent Feedback Form as an Audio Computer-Assisted Self-Interviewing (ACASI) instrument. In addition, Westat does not see a need to add more options to Q4.

Response Burden. Westat conducted data analysis to examine response rates to various tasks, the total number of food events and the total number of food items reported in the diary at the household level or person level, and the length of various interviews. One notable finding is that large households with at least four people have significantly lower response rates to the Initial Interview and the Final Interview and significantly lower rate to participate in all three telephone calls. Although large households reported significantly more food events and more food items at the household level, they reported significantly fewer per-person food events and food items than smaller households. In addition, across all interviews, households with four or more people took significantly longer time to complete the interviews than smaller households. This trend confirms that large households have a more burdensome reporting task than smaller households as indicated by the longer interview time. The higher level of burden faced by large households led to lower response rates to all tasks. We propose several potential strategies to encourage participation from large households such as assigning best interviewers, providing additional incentives, and providing more timely feedback and reminders.

Summary. Looking ahead at the next main data collection, Westat has two major recommendations to reiterate here. First, Westat firmly believes that ERS should consider fully exploiting the computerization of all instruments and forms to maximize the functionalities of computerized instruments and forms. Second, Westat strongly recommends that ERS explore an alternative food reporting system that can be accessed from multiple interfaces and devices to replace paper food books and forms. Westat suggests that ERS be sensitive about the level of reporting burden on large households and be proactive in engaging their continued participation throughout the study week.

1 Introduction

The 2012 National Household Food Acquisition and Purchase Survey (hereafter referred to as "FoodAPS-1") gathered detailed information about household food acquisitions from April 2012 to mid-January 2013. The survey was sponsored by the U.S. Department of Agriculture (USDA) and developed and fielded by Mathematica Policy Research (Mathematica). The nationally representative sample consisted of nearly 5,000 households that completed the FoodAPS-1 final interview. FoodAPS collects comprehensive data on American households' food acquisition, factors influencing food choices, and household well-being. In 2015, the Economic Research Service (ERS) contracted with Westat to conduct an independent assessment of the quality of the FoodAPS-1 sample design, instrumentation, data collection procedures, and resulting data. This report is part of a series of five reports that constitute that assessment.

The survey includes two in-person visits (one at the beginning of the reference week and one at the end of the reference week); three telephone interviews intended to be initiated by sample households; and five paper-pencil forms (Primary Respondent Food Form, Adult/Youth Food Form, Meal and Snacks Form, Income Worksheet, and Feedback Form) to be completed by sample households throughout the study week. The overall weighted response rate was 41.5 percent.

The ERS of USDA awarded a contract to Westat to assess all instruments and study materials employed in FoodAPS-1 against best practices in the survey field and the latest empirical evidence in the survey literature. To carry out a comprehensive assessment of all instruments and study materials, Westat first thoroughly reviewed documentations provided by ERS, including codebooks and the Data Quality memo series. Westat also collaborated with ERS to identify main concerns with each instrument and form. Westat then conducted expert reviews on various instruments used in FoodAPS-1 as well as the design features of FoodAPS-1. Secondary data analyses are also conducted to assess instrument designs and response burden. Recommendations are made for ERS' information based on the reviews. However, Westat strongly recommends that ERS evaluate the recommendations and test the selected recommendations before implementing them.

Instead of reviewing all questions in all instruments, the assessment is conducted with ERS' main concerns as the focus. The report is organized by instruments and study materials. Section 2 reviews the paper screener form. Sections 3 and 4 assess the Computer-Assisted Personal Interviewing

(CAPI) instrument used for the Initial and Final In-Person Interviews. The food books, Meals and Snacks Form, and the Income Worksheet are discussed in Chapter 5. Chapter 6 reviews the Telephone Interview Instrument while the Respondent Feedback Form is reviewed in Chapter 7. Chapter 8 assesses response burden and response rates, and Chapter 9 reviews the use of incentives. Conclusions and recommendations are offered in Chapter 10.

Review of Screener Form 2

2.1 Overall Issues with Screener Form

FoodAPS-1 employed a paper-pencil form for the screening interview. The overall screener response rate (SRR) was 72.2 percent. A review of ERS documentation revealed several issues with the paper screener form. First, paper screeners were missing for 2 percent of cases that were screened. Possible causes include errors in recording personal IDs (MPRIDs) on the paper screeners, failure of interviewers to return all materials, and incorrect assignment of screener status to complete where the case was not actually screened. Second, the screener data had duplicate records for some MPRIDs because field staff made errors in transcribing the MPRID onto the paper screener. A small number of screeners (n=33) were discarded because multiple records exist for the same MPRIDs with different household information (different household size and screener respondent name or phone). Third, inconsistencies exist within the instrument, especially in the presence of skip patterns. For instance, about 0.3 percent to 1.5 percent of cases are missing information on the gate questions (e.g., Q4, Q5, Q6, Q8), but the subsequent questions are non-missing (e.g., Q4a, Q5a, Q6a, Q8a). Fourth, interviewers sometimes either failed to fill the household size and income category box or they filled the boxes but not the questions. The frequency for this type of error ranged from 0.8 percent to 6.4 percent.

Survey literature has demonstrated that, compared to paper-pencil forms, computerized instruments have the benefits of reducing missing data by automating the process of determining which question comes next (Tourangeau, Rips, and Rasinski, 2000). As a result, a computerized screener form would have eliminated interviewers' failure to follow skip instructions correctly. In addition, computerized instruments can prefill information such as sampled addresses, MPRIDs, and date and time of data collection, again reducing the probability that an MPRID is entered or linked to a sample address. More importantly, a computerized screener form could program various edit checks to check for out-of-range values, inconsistency values, and outlying values, and it could afford the opportunity for interviewers to resolve these issues on the spot. Consequently, Westat recommends a computerized instrument for the screener interview for the next main data collection (we refer to it as "FoodAPS-2"). Westat understands that the expected eligibility rate is low and that interviewers are expected to find only about two eligible households agreeing to the survey each week, but we do

not foresee that a computerized screener form would incur an additional burden for interviewers other than bringing a laptop or tablet computer to the sampled households.

2.2 Review of Q1 to Q5b

Westat reviewed the first five questions of the screener. Q1 (in the first version of the screener only), Q2, and Q2b and Q2c (both are only in the second version of the screener only) are part of the survey introduction and process of obtaining permission. Q3 verifies the sample address and the interviewer terminates the screening interview if the respondent's address does not match the sample address. The Q5 series (Q5, Q5a, Q5b) establishes whether or not there is an additional housing unit at the sample address. The answer to Q5b (Do the occupants of the additional units or living quarters live separately from the people in your household?) determines whether or not the additional unit should be considered part of the household and be counted in household size. These questions are necessary for sampling purpose and should be kept where they are now.

The Q4 series (Q4 and Q4a) captures the respondents' mailing address. According to ERS, these questions are asked upfront in order to identify and confirm the sampled address, which will be used for comparison to addresses obtained at Q5a. Given this intended use, Westat agrees that these questions should be kept where they are now.

2.3 Review of Q6 and Q6a

The Q6 series (Q6 and Q6a) checks whether the sampled housing unit is seasonable or not. Seasonable houses (i.e., housing unit that are occupied for less than 6 months during the year) are not eligible for the main study. These two questions are important to establish the eligibility of sampled housing units and are, therefore, recommended to be asked upfront.

2.4 Review of Q7-Q8a

Q7, Q8, and Q8a are jointly used to capture household size. Household size is then used to determine the income level that qualifies households for program participation. Westat understands that the definition of "household" used in FoodAPS is broader than used in most surveys, and

Westat makes suggestions only to the wordings of the questions measuring "household" and not to the definition of "household."

Q7 asks for the number of people living at the sample address. Instructions are provided after the question proper to remind respondents to both include and exclude certain types of people, which makes Q7 long and complicated. The first instruction ("Do not forget to include babies, small children, and non-relatives who live here.") is a double negative, which is cognitively harder to understand than the alternative explicitly stating who to include (e.g., "Please include babies, small children, and non-relatives who live here.").

Four major government-sponsored surveys – the Current Population Survey (CPS), the National Health Interview Survey (NHIS), the National Health and Nutrition Examination Survey (NHANES), and the Consumer Expenditure Interview Survey (CE) – adopt a different approach, which starts with the question proper ("How many people live in your household?") and then follow up with one or more questions to determine whether respondents include the right people. For instance, CE follows up with one question ("Is there anyone else living or staying here now – any babies, small children, non-relatives, or anyone else?"), whereas CPS, NHIS, and NHANES follow up with three probes ("Have I missed any babies or small children?"; "Have I missed anyone who usually lives here but is away now – traveling, at school, or in a hospital?"; "Have I missed any lodgers, boarders, or persons you employ who live here?"). Westat recommends that ERS implement a three-part question, as presented in Table 2-1. Q7a asks directly the number of people usually living in the household, and Q7b probes about people to be included. The numerical answers to Q7c override the numerical answers to Q7a as the correct household size.

The Q8 series (Q8 and Q8a) checks meal-sharing status as part of the household definition. Westat does not see a problem with the wordings of Q8 and Q8a. However, we do recommend ERS computerize the screener form to resolve the issue of inconsistencies between Q8 and Q8a as observed in FoodAPS-1.

Table 2-1. Recommended question wordings for Q7

Q7a. Including yourself, how many people usually live in your household? Do not include people living away at school. _____NUMBER DO NOT KNOW REFUSED **Q7b. Is there anyone else living here, including babies, small children, non-relatives who live here, or anyone who usually lives here but is temporarily away for reasons such as vacation, traveling for work, or in the hospital?** YES→GO TO Q7c NO→GO TO Q8 DO NOT KNOW→GO TO Q8 REFUSED→GO TO Q8 **Q7c. Now including any babies, small children, non-relatives, or anyone who is temporarily away, what is the correct total number of people living here?** _____NUMBER DO NOT KNOW REFUSED

2.5 Review of Income Questions (Q9 to Q11)

Question Q9 has no reference period and does not tell respondents whether they should think about only regular income source or all sources (including a one-time payment). Both contribute to discrepancies between Q9 and the F1 series in the Final Interview asking about income sources. Westat suggests at least changing the wording of Q9 to include a reference period (e.g., "in the last 12 months") and some text to encourage respondents to report irregular income sources. Below are suggested question wordings for Q9:

> From now on when we refer to your household, we mean the [FILL HH SIZE] people that live together and share food.
>
> Next are questions about your household's income in the last 12 months, that is, since {{MONTH},1, {YEAR}}. When we say "income" we mean earnings from work, unemployment, welfare, child support, retirement income, disability income, investment income, and any type of income, even if you do not get it regularly. Please look at this card [SHOW INCOME SOURCES HAND CARD] and tell me which types of income were received by people in your household in the last 12 months.

> [IF NECESSARY: SNAP BENEFITS ARE NOT COUNTED AS INCOME.]

We recommend changing the wording of Q10 to be consistent with Q9:

> Thinking about your household's income from [LIST INCOME SOURCES IN Q.9] in the last 12 months, which group (A, B or C) corresponds to your household total income in the last 12 months, before taxes? SHOW HAND CARD FOR THE HOUSEHOLD SIZE.

A computerized instrument can automatically fill in income threshold values based on household size for Q10a and Q10b, eliminating the possibility of interviewers filling in the wrong income value.

A second issue with the income questions in the Screener is that total income reported at the screening interview was lower than total income obtained from the Final Interview. As a result, some households originally identified and recruited as non-SNAP households with low income relative to the Federal poverty guideline ended up being categorized as non-SNAP households in higher income groups. Research shows that the place at which one begins the initial income cutoffs makes a difference in the results (Monsees and Massey, 1997; Hill, 1999). Using FoodAPS-1 data, Westat examined the median difference between what is reported in the final interview for those who move into a higher income group in comparison to the cutoff for the income that they reported on the screener. Westat found that the median difference was between $5,000 and $10,000 depending on the size of the household. Altering the income cutoffs by this much would probably not be feasible and would increase reporting errors in the opposite direction. Westat also found that household size is significantly related to misclassification with larger households being more likely to be misclassified. Therefore, Westat suggests that ERS train interviewers extensively on the income questions (both A9 and A10 at the Screener Interview and the F-series at the Final Interview), allow interviewers to adopt a more conversational interviewing style when asking income questions, and encourage interviewers to prompt respondents from households with more than one adult to include income from every household member.

ERS could also consider dropping Q10 and just keeping Q10a and Q10b in the instrument. In other words, everyone is asked Q10a first on whether the total household income in the last 12 months is more or less than 100 percent of the Federal poverty guideline. For those who answered "more," Q10b follows up by asking if the total income is more or less than 185 percent of the Federal poverty guideline. This way, ERS reduces the burden for 9 percent of respondents who provided missing data to Q10 in FoodAPS-1. ERS is concerned that using Q10a and Q10b could have a

negative impact on response rates to the income questions. Q10a and Q10b employ unfolding brackets. Unfolding brackets are considered an effective strategy to reduce missing data to income questions (see Juster and Smith, 1997; Yan, Curtin, and Jans, 2010). The unfolding brackets approach simplifies the response task (only a simple yes or no answer is needed) and reduces the perceived sensitivity by conveying the message that an exact amount is not necessary. Consequently, the unfolding brackets approach is successful at reducing item missing rate. Although we are not aware of empirical work comparing item missing rates to questions such as Q10 and the unfolding brackets approach, we do not foresee a reason why the unfolding brackets approach would increase item missing rate in comparison to Q10.

Review of the Initial Interview Instrument 3

3.1 Overall Issues with the Initial Interview Instrument

The Initial Interview is conducted via CAPI. Unlike the screener form, the Initial Interview instrument is computerized. However, the full potential of a computerized instrument was not taken advantage of. For instance, about 136 households were identified as having errors that had to be reviewed manually. Examples included answering "Child" to A5 (the question asking the relationship of a household member to the respondent) but with an age greater than the respondent age. Edits can be added to the instrument so that CAPI will check for inconsistencies (as mentioned earlier), outlying values, out-of-range values, and so on. Westat reviewed the Initial Interview instrument and suggested out-of-range edits to be added on the following items: A6, B7a, B11a, B11b, B14c, and C14.

Westat further recommends passing information from the screener instrument (once it is computerized) to the Initial Interview to set up the context for the Initial Interview. For instance, the screener asks about the number of people usually living at the sampled address. This piece of information can be passed on to the Initial Interview for two potential uses. First, household size can be used in the question wording at A1 so that interviewers can start by asking for the names of those people at A1. Second, this piece of information can also be used to check against the number of household members listed at A1a. Respondents will reconcile the discrepancy if there is one.

3.2 Review of Section A

3.2.1 Review of Questions A1 to A3c3

The purpose of Questions A1 to A3c3 is to make a roster of people who live or stay at the sampled address. Westat reviewed the wordings of these questions and had three concerns.

First, we find the instruction before A1 (below) confusing in terms of reference period.

> "We will begin the interview with questions about who lives here. I need to make a list of all the people who are living or staying here at this address over the next week. Be sure to include: People who stay here only some of the time; non-relatives who live here; and any babies and small children. Please mention someone even if you're not sure they should be included. I understand that you may have already provided some of this information, but I need to read the whole series of questions."

It is not clear from the instruction whether respondents should be thinking about people who usually live at the sample address or people who will be living at the sample address over the next week (i.e., the data collection week). If the intended purpose of the instruction is to make a list of people who will be living or staying at the address during the diary week, then the following Question A1_CK ("Does anyone else live or stay here?) needs to be revised to be consistent with the "next week" timeframe. By contrast, if the purpose of the first two questions (A1 and A1_CK) is to capture people who usually live or stay at this sample address, then the sentence about making a list of all the people who are living or staying here at this address over the next week should be removed.

Second, Westat is concerned about the length of the A3 series. Westat understands that Questions A3 to A3c3 probe for people who tend to be overlooked when enumerating household members. Questions A3b2 to A3b4 ask for the *number* of lodgers, boarders, and persons that respondents employ who live at the sample address. Questions A3c1 and A3c3 ask for the *number* of people who have another place to live but stay at the sampled address often, or have some space or a room at the sampled address. Westat recommends that ERS re-evaluate the purpose and the potential use of these five questions asking for numbers. If these numbers do not play a major analytic role, ERS should consider dropping these items from the Initial Interview to reduce response burden.

Third, Westat recommends moving the A3 series before the A2 series. In other words, the recommended order is A1 and A1_CK (people living or staying at the sampled address), A3 to A3c3 (people who are likely to be overlooked), and A2 to A2d (establishing main residence). In this proposed new order, ERS will obtain a complete list of people living or staying at the address, their main residence, and their residence in the diary week.

FoodAPS-1 documentation does not seem to indicate that item nonresponse to questions A1 to A3c3 to be an issue. But if there were a concern about respondents unwilling to provide a full name

to the A1 name grid, literature shows that asking for nicknames or initials (instead of full names) reduces underreport of household members (Tourangeau, Shapiro, Kearney, and Ernst, 1997).

3.2.1.1 Review of Question Wordings to Evaluate Inclusion of People Contributing to Income

ERS is concerned that household members who contribute income to the household but are out of the house for extended periods of time (such as deployed service men and women) and, thus, are not captured in the Initial Interview. And the income from these household members is not captured in the Final Interview since they are not listed in the household roster. Westat reviewed the wordings of Questions A1 to A3c3 and found that ERS attempted to include deployed service men and women through the instructions before Question A1 and Question A3c. The instruction before A1 encourages respondents to include "people who stay here only some of the time." Question A3c probes for people "who may have another place to live, but who stay here often or have some space or a room here." However, the instruction and the question may not be sufficient for ERS to capture *all* deployed service men and women. Especially for A3c3, only those who will be staying at the residence the next week are added to the household roster. As a result, Westat recommends that ERS revise the instruction before A1 to remind respondents to report all people who have a room or stay at the sample address and are currently away either temporarily or for an extended period of time (e.g., deployed service men and women, college students, children at overnight summer camps etc.). In addition, Westat recommends that ERS revise Question A3c3 to add all people who stay at the sample address or have a room at the sample address to the household roster. According to ERS, the ultimate goals of this section are to identify all people whose income contributes to food expenditure decisions and to accurately capture the number of people eating food during the data collection week. Westat further recommends adding one more question after A2d2 to check whether or not household members listed on the household roster contribute financially to the household.

3.2.1.2 Review of Question Wording to Include Stay-Over Guests and Meal Guests

Questions A3 to A3c3 might have been intended to include stay-over guests and meal guests. However, these two types of people are not explicitly asked about during the Initial Interview. Westat recommends that two new questions be added to the Initial Interview to specifically ask about stay-over guests and meal guests for the diary week so that these two types of people can be accurately accounted for in the Meals and Snacks Form and food acquisition reporting. Of course, it

is very likely that respondents do not know the total number of guests ahead of time or that unexpected guests may show up during the data collection week. In that case, respondents would be instructed to record, in the food books and on the Meals and Snacks Form, guests that join the household at any time during the data collection week and to report food acquisitions by and for these guests during the week.

3.2.2 Other Issues in Section A

The A9 series evaluates the ability of the Initial Interview respondents to report food acquisitions during the diary week. Westat understands the importance of getting this piece of information. However, the A9 series, in its current order, interrupts the flow of the questionnaire and respondents have to move from person-level reporting at A4 to A8a, to household-level reporting at A9 to A9e3a, then back to person-level reporting at A10 again. According to ERS, the rationale for including these questions here after A8a is to identify households where no one meets the survey requirements and then to immediately stop the interview for those households. If the sole purpose is to spare households where no one meets the survey requirements for the interview, it seems that the best place to ask these questions is at the screener after the main food shopper is identified. However, given that no household was dropped after A9 series in FoodAPS-1, Westat recommends that the A9 series be moved to the end of the Initial Interview so that interviewers can immediately speak with the right person(s) to train them about the food acquisition reporting tasks.

3.3 Review of Section B

3.3.1 Review of Questions B1 to B4

Questions Q1 to B3c measure SNAP participation (current participation or participation during the last 12 months) and SNAP benefit amount. Q1 is intended to capture whether or not households currently receive SNAP. For those who responded "Yes," they are followed up with B1, which asks about the date of their last receipt. For those who answered "No," "Don't Know," or "Refused" to Q1, they are followed up with B3, which asks whether or not they ever received SNAP. B3a is intended to capture the date of last receipt. One problem with Q1 is that it does not specify a reference period in the question wording. As a result, ERS found, from FoodAPS-1, two issues related to self-reported dates of SNAP receipts. Some respondents answered "Yes" to the current

receipt question (Q1) but reported dates of last SNAP receipt (at B1) older than 30 days ago. Similarly, some respondents reported "No," "Don't Know," or "Refused" at the current receipt question (Q1) and then reported a date within the last 30 days (at B3b). ERS suggests a sequence that first asks if anybody in a sampled household received SNAP benefits in the last 12 months. For those who responded "Yes" to this question, they will be asked who in the household received the benefits and the date they last received the benefits. If the last receipt date is within the last 30 days of the interview day, respondents will be asked to confirm if they are current SNAP participants and whether they expect to receive more benefits. Those who expect to receive more benefits will be asked when they expect to receive the extra benefits. Westat welcomed the suggestion and saw the potential of this sequence to reduce the reporting issues with the last receipt dates as observed in FoodAPS-1. However, Westat does think that the full sequence should be tested in advance of fielding it in FoodAPS-2.

Alternatively, a reference period could be specified in Q1 (e.g., Since [MONTH] 1, [YEAR], (Do you/Does anyone in your household) currently receive benefits from the SNAP program?") to specifically ask about SNAP participation in the last month and to reduce the likelihood that respondents reported early dates for current participation and recent dates for prior participation.

Another problem with questions Q1 to B3c is that they measure SNAP participation at the household level; that is, ERS only learns whether or not someone in the household receives the SNAP benefit from these questions. ERS does not know which specific individuals in the household receive SNAP benefits and who belongs to which SNAP unit. This is true especially for large households with complex living arrangements where multiple SNAP units may live at the same address. Additional questions will be needed to fully capture SNAP units as defined by SNAP regulations. However, the definition of SNAP units might not be fully understood by respondents and is hard to convey to respondents. Westat recommends adding one additional question after B1 that asks respondents to select from the household roster (constructed at Section A) specific individuals in the household who receive SNAP benefits. This additional question enables ERS to know who in the household receives SNAP benefits and to construct SNAP units using these individuals' characteristics (such as age, relationship, and income) at the analysis stage. Although the limitation of this approach is that ERS would not be able to construct separate benefit amounts for each SNAP unit in the case of multiple SNAP units living within one household. ERS may not consider it critical to have separate benefit amounts for each SNAP unit.

Westat also suggests that ERS evaluate the definition of SNAP units and eligibility criteria to determine what additional pieces of information may need to be included in the instrument to fully identify SNAP units and benefit amounts.

3.3.2 Review of B14 Series

Question B14 is a filter question asking if anyone in the household is receiving benefits from the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), and B14a asks respondents to select from the household roster WIC benefit recipients. In theory, these two questions together should identify all household members receiving WIC. However, it is possible that respondents fail to select all WIC recipients at B14a. One method to reduce possible underreporting at B14a is to have one follow-up question asking respondents to verify that any child under the age of 6 in the household not selected at B14a is, indeed, not receiving WIC benefits. Furthermore, B14b should be asked for all female household members aged 12 or older, regardless of their pregnancy status. For each member who is reported as "receiving benefits for a child or for both self and child," CAPI should check to see if a child is selected at B14a and, if yes, verify that the selected child(ren) at B14a is/are the one/ones receiving WIC benefits. When CAPI cannot find a child at B14a, interviewers should probe to ask which child the female household member is receiving WIC benefits for.

3.4 Possibility to Move Questions to a Web Interview

Section C asks respondents where they usually get food. There are questions asking for specific store names and store addresses. Westat thinks that this section is a candidate to be moved to a web interview if needed. A web version of Section C has the potential to ease the reporting task, especially the task of reporting names of stores visited and the location of those stores. For instance, Google Maps can be embedded in the web instrument to allow respondents to zoom in on areas where they live and shop. This is feasible if ERS considers the possibility of implementing an app-based or a web-based diary to capture food acquisition for FoodAPS-2. The programming and fielding cost of adding more survey questions to the app-based or web-based food reporting system should be marginal.

Review of Final Interview Instrument 4

4.1 Overall Issues with Final Interview Instrument

Same as with the Initial Interview, the Final Interview is also conducted via CAPI. The full potential from the computerization of the instrument should be taken advantage of to increase the instrument's ability to reduce reporting error. Edit checks can be programmed into the instrument so that data anomalies can be identified and verified with respondents during survey administration. For instance, the expected range for item A1 (number of times respondents or households preparing food for dinner or supper at home during the past 7 days) is 0 to 7. Logic checks could be programmed into the CAPI instrument so that any answers provided that are outside this range will be prompted for further verification. Westat reviewed the Final Interview instrument and recommends out-of-range or outlying values checks to be programmed into the instrument for the following items: A1, A2 (number of meals family sitting down and eating together in the past 7 days), A3a (number of days guests coming for a meal or snack in the past week), A3d (number of guests on a particular day), D3 (height), D4 (weight), and I1 (month and year of moving to current residence).

Two questions (E5a and E9a) are part of the food security questions. Both ask the number of days certain things happened in the last 30 days. ERS is concerned that implementing edit checks on these two items would introduce a change that could disrupt comparability across surveys. Westat understands ERS' concern and suggests that at least soft edit checks be programmed to alert interviewers when a respondent provides a number larger than 30 and to flag those cases for further review and post-survey editing. An experiment could be implemented in a pretest where some respondents are randomly assigned to receive the food security questions without edit checks and the rest to receive the food security questions with edit checks to examine the impact of including edit checks on the answers to the food security questions.

Questions on income and expenditure (e.g., F2, F3, F4, F5, F5c1, F6, F8b, G1a, G2a, G3a, G4a, G5a, G6a, G7a, G8a, G9a, G10a, G11a, G11d, G12a, G13a, and G14a) can also benefit from programmed edit checks to reduce data anomaly. ERS suggests using different edit checks based on ZIP codes or Census tracts. Westat is not aware of empirical studies or surveys using different edit checks based on ZIP codes or Census tracts, but sees the advantage of doing so. The challenge is to

locate or construct possible ranges at a reasonable geographic level. Westat doubts that expenditure estimates are available at the ZIP codes or Census tracts. The Consumer Expenditure Interview Survey (CE), for instance, only includes state and (some) primary sampling units (PSUs) in their public use dataset. The 5-year or 3-year PUMS data from the American Community Survey (ACS) might provide income ranges at the Census tracts level, though. Westat suggests that ERS explore the right level of geography and evaluate the possibility of using different edit checks at that geographic level.

4.2 Review of Body Measures

Question D3 measures the height of household members in either English or metric units. D3 has item nonresponse of 1.4 percent. Weight is measured through Question D4 and has item nonresponse of 3.2 percent. The two follow-up questions (D4_Alt1 and D4_Alt2) provide additional information on whether a household member is overweight or obese when respondents failed to provide an answer to D3. These two follow-up questions reduce the level of item nonresponse for indicators of overweight and obesity constructed later during analysis.

During analysis, reported height and weight are used to calculate body mass index (BMI). According to the ERS documentation, the CAPI instrument did not check for outlying values in the height and weight data, and these outlying values sometimes produced questionable BMIs. For instance, the BMI is less than 15 (severely underweight) for 27 individuals and over 40 (very severely overweight) for 532 individuals. Westat offers two recommendations as an attempt to reduce the chance of anomalies in height, weight, and BMI data. First, edit checks could be programmed into the CAPI instrument for D3 and D4 separately so that any values outside acceptable ranges could be verified with respondents during survey administration. Second, the CAPI instrument could calculate BMI once both height and weight were available for a household member during survey administration. If the calculated BMI was out of an acceptable range, interviewers could verify both height and weight data before moving to the next questions. These two steps should help reduce the amount of reporting error in body measures. ERS wondered whether the edit ranges can be based on a person's age. From a CAPI programming perspective, Westat thinks it is feasible. However, the challenge again is to locate or to construct the acceptable ranges based on age. As long as those ranges are available, programming them into the CAPI instrument would not be hard.

4.3 Review of Income Questions

Member-level income questions (for persons age 16 and older) have item nonresponse between 2 percent and 5.4 percent. There are also significant missing data for the frequency of income questions. In addition, some respondents may have reported zero income as a "passive refusal" to the income questions. The questions on the Final Interview were designed to reference the Income Worksheet and referred to the income categories as they appeared on the worksheet. However, given that at least 40 percent of respondents did not complete the Income Worksheet, Westat is concerned that the current question wordings might pose challenges for respondents. Westat reviewed the question wordings of the F1 series (F1 to F8b) and noted several possible comprehension issues.

Question F1 asks if respondents had any income last month. The question is vague in that it does not explain what constitutes as income and whether or not income here includes irregular earnings (e.g., one-time earning from babysitting a neighbor's children). F1b is added as an edit check to reduce the prevalence of reporting "no income" to F1. Westat recommends that ERS replace F1 with F1b; F1b clarifies the concept of income and encourages respondents to include irregular income.

The F2 series asks about earnings from work. But it is not clear whether tips, commissions, and bonuses count as earnings from work. Also, the question implies regular earnings from work. Westat recommends that ERS clarify the meanings of "earnings from work" (by presenting respondents with a showcard that lists all types of earnings to be considered) and add text to remind respondents to include one-time work.

The F4 series measures income from welfare, child support, and alimony. The question does not state explicitly whether respondents should include, for instance, SNAP benefits. In addition, answers to the F4 series cannot allow ERS to calculate income from SNAP, WIC, Temporary Assistance for Needy Families (TANF), or other assistance programs. Westat recommends (1) separating welfare income from income from the other two sources (child support and alimony) and (2) presenting a showcard to list all types of welfare income to be considered. If ERS is interested in the ability to obtain income from each type of welfare program (e.g., SNAP, WIC, TANF), Westat recommends that ERS ask respondents about income received from each type separately. ERS can also bring in responses to Q1 and B14 of the Initial Interview to reconcile with

answers to F4 about receiving income from welfare. Discrepancies could be resolved during survey administration.

Questions F5 to F5b measure income from retirement and disability. The question wordings for F5 do not tell respondents what to include and what not to include (e.g., workers' compensation is listed as an income source at F5b, but F5 does not mention it at all). Westat suggests that ERS show the income sources on a showcard to respondents, ask respondents to select (from the showcard) sources from which they received income during last month, and then ask respondents to report the amount of income received across the sources they identified.

Questions F5c to F5c3 measure income from investment. Again, Westat recommends showing all investment income sources on a showcard and asking respondents to select the income sources first before asking about the total amount of income received across the identified income sources.

The F6 series measures income from other sources that have not been asked in the interview. Westat recommends showing all types of income sources on a showcard and asking respondents to first report the other types of sources of income before the income amount.

The F7 and F8 series serves as an edit check to verify income sources and income amount reported. Westat suggests that ERS make use of the income source information collected at the screener interview to probe those who reported No/Don't Know/Refused at F1 or F1b. ERS could remind this group of respondents the sources from which the household received income last year (as responded during the screener) and ask them to confirm whether anyone from the household did or did not receive any income from those sources during the last month. When respondents refuse to confirm (through Don't Know or Refused answers), Westat recommends that ERS present a list of income categories to respondents and ask respondents to select one to further reduce overall item nonresponse to income data. The income categories can be used to impute an exact income amount before being used together with the collected exact income amount in the analysis and computation of weights (Juster and Smith, 1997). Alternatively, they can be combined with range categories computed from the exact income amount in the analysis and computation of weights.

An additional procedure to reduce item nonresponse to income data is to provide a list of income categories to respondents who answered Don't Know or Refused to each income source. This procedure will reduce the overall level of item nonresponse for two reasons (Juster and Smith, 1997; Yan et al., 2010). First, it is cognitively easier to choose an income category than to report a dollar amount. Second, questions asking respondents to choose one income category are perceived to be

less sensitive. But this approach does add to the overall length of the interview. It is recommended to evaluate the importance of income questions to the objectives and make a trade-off decision.

The CAPI instrument can be programmed so that the reported income amount from different sources can be aggregated to member level and even to household level. Edit checks can be built into the CAPI instrument to check for outlying values or values outside acceptable ranges; amounts that are too low or too high can be verified with respondents during survey administration.

Lastly, ERS should consider another potentially promising technique for improving the measurement of income – the use of conversational interviewing techniques whereby the interviewer is better able to clarify potential conceptual problems with measuring income. A recent paper has shown that this approach may help to reduce bias in income reporting without significantly increasing interviewer variance (West, Conrad, Kreuter, and Mittereder, 2016). Westat believes that intense trainings for interviewers on the income section could be used to motivate a conversational interviewing approach on the screener and the final interview. Interviewers could also be provided additional information (such as Q by Qs, additional scripted probes to be used under different circumstances) to guide their administration of the income section.

4.4 Review of Expenditure Measures

Item nonresponse to expenditure items (Section G) ranges from less than 1 percent to 8.7 percent. A data anomaly is found on G14 (expense for adult care); one respondent reported $15,840 to G14 when the expected range is from $0 to $7,409, according to the Data Quality Memo.[1] Westat reviewed the question wordings of expenditure items and did not find issues with the question wordings that would affect respondents' understanding of the questions.

However, expenditure questions are cognitively hard to answer as most people do not remember the exact dollar amounts they spent. The CE – a national study providing comprehensive information on a wide variety of expenditures – faces the same challenge and has conducted research looking into respondents' use of records and the impact of using records on data quality. The current protocol for CE is to encourage respondents to use records whenever possible to answer expenditure items, and one study finds that the use of records is associated with improved data

[1] Prepared for ERS by Mathematica, dated March 20, 2013.

quality, particularly for certain types of expenditures, such as payments for utilities and health insurance (Edgar, 2010). A small-scale field test conducted for CE compared self-reported expenditure amounts to records and found that a little over half of the reported expenditure amounts (53%) were within 10 percent of the correct amount (Kopp, 2013). Respondents misreported their expenditure amounts by 36 percent on average (Kopp, 2013). Another pilot study examining the feasibility of asking respondents to collect and save records to be used in a future study shows that over 80 percent of the respondents thought that keeping records was easy, and they did not feel uncomfortable sharing all of their expenditure records (Sjoblom and Lee, 2012). Based on CE's research, Westat recommends that ERS instruct respondents at the end of the Initial Interview to collect records (e.g., utility bills, credit card statements, bank statements, etc.) to help respondents answer these expenditure items.

This set of items is also a candidate to move to a web interview if needed. A web version can remind respondents to use records, suggest types of records that respondents can use to answer different expenditure questions, and allow respondents enough time to retrieve records and to consult other household members to come up with the answers.

4.5 Placement of Food Security Questions

ERS is concerned that answers to the food security Questions E1 to E9 could be affected by respondents' participation in the week-long food acquisition recording. The week-long recording of food acquisition might have heightened the sense of food security (or insecurity) for respondents, biasing their answers. Westat agrees that a context effect on food security questions possibly exists, arising from being placed in the Final Interview after the week-long collection of food acquisition. However, placing the food security questions in the Initial Interview could have an impact on the reporting of food acquisition, too. Westat recommends that ERS conduct a methodological experiment to assess the placement effect. Specifically, the experiment randomly varies the placement of food security questions so that a half of the respondents are asked the food security items at the end of the Initial Interview before they start recording food acquisition and the other half receive the food security questions as part of the Final Interview after they are finished recording food acquisition. Due to the random assignment, any observed differences in answers to the food security items between the two groups can be attributed to the placement of the items.

4.6 Identifying TANF as a Source of Income

As discussed in Section 4.3, the current questions of the F4 series will not be able to identify TANF as a separate source of income. The F4 series lumps together welfare, child support, and alimony. As a result, the exiting set of questions do not distinguish income from welfare, child support, and alimony. They also do not distinguish income from specific types of welfare. As recommended in Section 4.3, types of welfare should be presented to respondents on a showcard, and respondents should be asked to select from the showcard the types of welfare programs from which they received income. Separate questions can be used as follow-up, asking the amount of income from the identified sources.

4.7 Identifying SNAP Benefits Receipt During Data Collection Week

The current questions do not allow ERS to identify whether an issuance of SNAP benefits was received during data collection week. Westat does not make suggestions on whether or not SNAP benefits should be treated as income because the answer depends on ERS' analytical goals. Westat finds that, even though F1 asks whether household members had any income in the last month and F4 asks amount received from welfare, child support, and alimony, answers to these questions do not tell directly whether household members receive SNAP benefits during the data collection week. There are two issues. First, last month could mean the last calendar month (which may or may not include the data collection week), or the last 30 days (which definitely includes the data collection week). Second, there is not clear direction to respondents at F1 and F4 whether respondents should consider SNAP benefits as income or not. Without conducting cognitive interviews to understand what respondents were thinking when answering these questions, it is hard to know how people perceived the last month to be and whether they included SNAP benefits in their answers to F1 and F4.

To capture receipt of SNAP benefits during the data collection week, new questions have to be added to the Initial Interview to ask whether respondents expected to receive SNAP benefits, and if yes, the expected benefits amount. At the same time, new questions have to be added to the Final Interview to ask whether respondents were issued SNAP benefits during the data collection week, the date and the size of issuance if yes, and whether the issued amount was the expected amount.

These new questions certainly add to the interview length and increase response burden. Westat recommends that ERS consider evaluating their objectives and then decide whether to add new questions capturing issuance of SNAP benefits during data collection week.

Review of Food Books, Meals and Snacks Form, and Income Worksheet

5

5.1 Reviewing the Meals and Snacks Form

The Meals and Snacks Form is a paper form to be completed by household members during the data collection week. The purpose of the form is to identify which meals and snacks were consumed by household members each day. Only 81 percent of household members responded to the form, meaning that the meals and snacks data are missing for 19 percent of household members.

The Meals and Snacks Form is edit-checked prior to key-entry. The edit-check identifies blank days (nothing checked) and distinguishes missing data (nothing checked for a person on a day) from meals that are simply "not checked." Westat recommends computerizing this form by making it an online form. A computerized Meals and Snacks Form would be able to solve the issues of blank days and missing data. For instance, error messages could be displayed back to respondents if a meal was not checked or when blank days happened.[2]

The data entry process identified two problems. First, the household member names from the Meals and Snacks Form do not always match with names entered in CAPI. The household member names are entered to the form by respondents with the intended purpose to engage them in the form. However, this resulted in names that were often spelled differently in CAPI (by field staff) and on the Meals and Snacks Form (by respondents). During data entry, household members were matched with CAPI data by assuming that the order of persons on the Meals and Snacks Form matched the order of persons listed for the CAPI household roster. Second, the days of the week were not written on many (n=931) Meals and Snacks Forms. Missing days are imputed by assuming that the first day was the same as the start of the data collection week. Westat believes that a computerized Meals and Snacks Form can solve these two problems. The names of household members (and even personal ID "PNUM") can be passed from the initial CAPI interview data to be populated on the computerized Meals and Snacks Form. Days of the week can also be automatically filled in. A computerized form saves the resources spent on manual checking the data and entering the data.

[2] Note that this discussion is not about the reporting of food acquired or purchased during the reporting week—the main objective of FoodAPS-1. The Meals and Snacks Form sought to identify which meals and snacks each household consumed on each day.

5.2 Reviewing the Income Worksheet

The Income Worksheet is an optional paper form provided to the primary respondents to be completed during the data collection week in preparation for the Final Interview. The primary respondents are instructed to list the names of all household members who had received income during the past month and to record the amount of income received in six income categories. The worksheet is intended to be used as a recall aid and is not collected by interviewers at the Final Interview.

During the Final Interview, interviewers indicated that only about 57 percent of respondents had completed the Income Worksheet in whole or in part. Since the income questions on the Final Interview were designed to reference the Income Worksheet and referred to the income categories as they appeared on the worksheet, about 43 percent of respondents might not have read and considered the definitions of the income categories as they appeared on the worksheet, reducing the utility of the worksheet to standardize respondents' understanding of income sources. Westat has proposed in Section 4.3 the use of showcards as a partial remedy to this issue.

The utility of the worksheet also depends on the extent to which respondents read through the texts below the table and remember to include income from those sources in their reports. As a self-administered form, it is not clear how many respondents actually read all the information on the form and how thoroughly they considered the definitions. To increase the utility of the worksheet to improve respondents' understanding of income and recall of income amount, Westat recommends that, at the end of the Initial Interview, interviewers take time to explain to the respondents each of the income sources listed on the worksheet and ask (and record) who in the household expects to receive from these income sources during the data collection week. This way, respondents' burden will be lessened and respondents' understanding of the income sources will be guaranteed.

5.3 Reviewing Food Books

ERS asks all household members age 11 years and older to track and report all food acquisitions through food books during the 1-week period. Three different types of food books are used. The primary respondent food book is intended for primary respondents, who are responsible for

recording all food-at-home (FAH) items acquired by any household members, recording own acquisition of food-away-from-home (FAFH) items, and recording FAFH items acquired by members younger than 11 years old.[3] The primary respondent food book includes Daily List pages, Red Pages (to record FAFH items), Blue Pages (to record FAH items), and Barcode Pages. The adult food book is intended for adult household members aged 19 or above, and it only contains Daily List pages and Red Pages. The adult household members are instructed to record FAFH items in their own food book (on Red Pages) and enter FAH items into the Blue Pages of the primary food book. The youth food book is for youths aged 11 or above and contains only Red Pages. The youths are instructed to record their acquisition of FAFH items on their own food books.

Red Pages and Blue Pages are reported by over 90 percent of households; 84 percent of households reported both Red and Blue Pages and less than 2 percent of households reported no acquisitions.

5.3.1 Collecting Food Acquisition Data

The documentation on FoodAPS-1 indicates that all Blue Pages are entered using a double-entry process whereby each Blue Page is key-entered twice by two different data entry coders. Inconsistent entries are resolved by the second coder.

Red Pages are captured in two ways. First, telephone interviewers enter the information into a Red Page data entry form in real-time when primary respondents call in for the Telephone Interviews or answer the outbound Telephone Interviews on days 2, 5, and 7 of the study week. Second, after food books are collected from respondents, Red Pages go through a "Red Page Review and Capture" process. All Red Pages from the physical food books are reviewed manually by reviewers and compared to the list of food acquisitions captured by the Telephone Interviews. If a Red Page is on the list captured by the Telephone Interviews, then the page is identified as reported by phone. Physical pages not matched to the telephone interview list are flagged for entry and sent to "Red Page Capture" where coders or telephone interviewers (between calls) enter Red Pages into the database through the Red Page Entry screen. Both pages require extensive manual review and manual entry, which takes time and resources.

[3] ERS refers to food and drinks brought into the home as food-at-home (FAH) and meals, snacks, and drinks obtained outside the home as food-away-from-home (FAFH).

In addition, several issues with data from the food books are identified in various documentations. For instance, there are duplicate Red Pages and Blue Pages that need to be reconciled and cleaned up. Duplicate food items are reported on both Red Pages and Blue Pages. To resolve these data duplication issues, decisions have to be made based on assumptions that may or may not be correct. Furthermore, there exist missing data (e.g., place name, address, dollar amount, reporting day of the week).

ERS is concerned that the characterization of FAFH and FAH is done poorly and that the use of two different food book pages may make things more confusing for respondents. It may be beneficial to consider an alternative approach that does not necessarily have to make a distinction between Blue Pages or Red Pages. For instance, under an electronic collection system, respondents could go to one interface and report all items they acquired during the study week. It may also be possible to keep the distinction between FAFH and FAH, but provide real-time guidance through images or text to convey the differences between the two types of events. An alternative could aim to reduce the confusion caused by different food book pages or different food item types.

An alternative electronic reporting system could be used to replace paper diaries for most households, which could have the ability to alleviate reporting burden and reduce the extent of post-survey data entry and processing. Westat is currently developing and testing such an approach for ERS in the ADCM Pilot Study.[4]

5.3.2 Identifying EBT Programs

ERS asked Westat to examine the instrument design features of the current food books and forms on the ability to better identify which program's electronic benefit transfer (EBT) cards are accessing. Westat reviewed the food books and forms and noted that use of EBT cards is asked on Red Pages and Blue Page. On Red Pages, "SNAP EBT" and "Other EBT" are offered as payment options. But the EBT amount is not asked separately for these two types of EBT cards. On Blue Pages, "SNAP EBT" and "TANF EBT" are offered as payment options. SNAP EBT amount and WIC amount are asked separately whereas TANF EBT amount is not asked.

[4] ADCM stands for Alternative Data Collection Method.

Based on the reviews of the food book pages, Westat believes that FoodAPS-1 is able to identify the use of SNAP EBT cards to pay for both FAH and FAFH items. FoodAPS-1 is also able to capture the amount of SNAP EBT paid for FAH items, but it may not be able to capture the SNAP EBT amount paid for the FAFH items, especially when multiple payments are checked off. In terms of EBT cards from other programs, FoodAPS-1 captures explicitly the use of TANF EBT cards to pay for FAH items, but not the amount. FoodAPS-1 also captures explicitly the amount of payment for FAH items coming from WIC, but it does not explicitly record the form of the WIC payment. The use of and the amount of TANF EBT and WIC EBT to pay for FAFH items are not explicitly recorded in FoodAPS-1. As ERS pointed out, WIC EBT is expanding its implementation, and it is necessary for future surveys to capture explicitly the type of WIC payments (WIC EBT vs. WIC paper checks).

In relation to the WIC payment, FoodAPS-1 does not allow the verification of WIC purchases of FAH items, and it is not clear to what extent WIC purchases are underreported. If it is possible to obtain a list of WIC foods from states ahead of time, the list can be used to identify WIC foods in the scanner data or receipts. Of course, this would require state-by-state programming and would need to be easy to update. ERS proposes another alternative where broad categories of WIC-permissible and non-permissible foods are programmed and used to identify WIC purchases. Westat recommends that ERS first explore the feasibility of obtaining a state-by-state list of WIC foods, evaluate the completeness and the accuracy of the list, examine the possibility of matching the list to the IRI database/scanned barcodes and receipts, and evaluate the feasibility of using broad WIC categories to identify WIC purchases.

5.3.3 Collection of Information on Bottle Deposits and Food Taxes

Large or negative price discrepancies between total itemized price and total paid are identified for both FAFH and FAH items in FoodAPS-1. One reason for the discrepancies is that FoodAPS-1 does not collect cost of non-food items, bottle deposit fees, and sales taxes. It is true that collection of these pieces of information (non-food items, bottle deposit fees, and sales taxes) would better help resolve discrepancies between total itemized price and total cost. However, Westat is concerned that collection of these pieces of information increases response burden and distracts respondents from better reporting food items, which is of more analytic interest to ERS.

ERS indicated that deposit fees and sales taxes are often considered part of the cost of food and that there may be no need to collect them. There is still a need, however, for FoodAPS-2 to identify

whether or not non-food items are included in a food acquisition event and the cost associated with these non-food items. Westat suggests that, for FoodAPS-2, respondents are explicitly asked to indicate on food books (or an alternative food reporting system) whether or not non-food items are present on a receipt. Westat also recommends that ERS explore additional ways to better match food items reported on food books to food items listed on receipts. For instance, respondents could be instructed to annotate receipts and provide the annotated numbers along with (or instead of) item descriptions and item prices on food books. Respondents could be also instructed to mark non-food items on the receipt so that the cost of non-food items can be calculated at the back end during post-survey processing.

5.3.4 Collection of Item Descriptions and Prices for FAFH Purchases

Red Pages collect the total paid for each FAFH purchase and the amount of tips. The instruction is for respondents to record FAFH item description, size or amount, quantity, and price when receipts do not list each food item or when receipts are not available. Information on Red Pages is then used as a memory aid for the telephone interviews. Data from these telephone interviews are the primary sources of information for FAFH.

ERS is interested in collecting all FAFH item descriptions and prices regardless of whether they are listed on the receipts or not. There are at least two advantages of collecting all item descriptions and prices. First, the collection of these additional pieces of information enables calculation of food items' contribution to a Healthy Eating Index (HEI) and enables ERS to obtain the product's nutrition characteristics. Second, the collection of the additional information also makes it easier to identify and remove potential duplicated reporting across Red Pages. The main disadvantage is, again, the increased reporting burden associated with entering from receipts food item descriptions and prices, and the potential negative impact of increased response burden on data quality and continued participation in keeping the food books. The challenge is to come up with a method to best collect all FAFH that balances burden and data quality.

As mentioned in Section 5.3.1, Westat recommends that ERS explore an alternative reporting system that would be equipped with features to reduce burden of reporting FAFH descriptions and prices. For instance, the alternative system could preload various pieces of information (such as dates, household member names, place names and addresses), and could use drop-down menus or autofill's as one types or looks up an item description. In addition, the alternative system could have

the capability to accept photos and memos that can be used by respondents to help with recall of FAFH and by ERS to help with post-survey editing and checking.

5.3.5 Collection of Price and Descriptions of Private Label and Other Unrecognizable UPCs

FoodAPS-1 asks respondents to scan the Universal Product Code (UPC) on every FAH item. If an item does not have a UPC code, respondents are instructed to look for a picture of the item in the Primary Respondent Food Book and scan the barcode next to the picture. If the item has no barcode and is not in the food book, respondents are instructed to enter an item description, size or amount, and price for that item in the Blue Pages. Scanned UPCs are then matched with a UPC data dictionary to obtain food item names. Items without a barcode or items whose UPC cannot be matched with the UPC data dictionary present challenges for standardizing item descriptions, and linking items to food categories and other extant information. ERS is considering collecting item description for private labels or UPCs that cannot be matched with the UPC data dictionary. Westat thinks it is a good idea to collect item description information for private labels. However, if ERS continues using paper food books, the scanned UPCs are matched to the UPC data dictionary only after data collection is over for the households. As a result, sampled households do not get notified in time whether a scanned UPC is able to match with the UPC data dictionary. An alternative reporting system that could provide timely feedback to respondents on the matching status of a UPC is desirable. After UPCs are scanned and uploaded to the alternative system, they could be matched to the UPC data dictionary and the matching status could be immediately provided back to respondents. In case of unrecognized UPCs, respondents will be asked to provide a description of the item. This type of timely feedback of matching status of UPCs is beneficial to data quality and reduces the extent of post-survey processing.'

5.4 Evaluating the Performance of the Scanner

5.4.1 How Well Did the Scanner Work?

FoodAPS-1 instructs respondents to scan the UPC codes on food items or the food pictures in the Primary Respondent Food book if they are present. The purpose of using a scanner to scan UPC codes is to reduce reporting burden on the respondents' part and to standardize item descriptions.

The scanned barcodes are later matched to a UPC data dictionary for product names. The UPC data dictionary combines information from Nielsen, Gladson, and Gregg London dictionaries into a single unduplicated list of barcodes consisting primarily of Nielsen data. Barcodes are also matched to the IRI dictionary after it was made available. Matched barcodes standardize the item descriptions and can also be used to look up nutrition values.

According to the End of Survey Memo, among the 193,693 FAH items reported by households, about 68.5 percent of items had scanned barcodes (either from the product or the Blue Pages) before ERS provided the IRI dictionary. Over 75 percent of the items with scanned barcodes (about 52.4% of total food items reported) were able to be matched to the UPC data dictionary for product names, which means that burden was successfully reduced for both respondents and post-survey processing personnel. The rest of the scanned barcodes could not be matched to the UPC data dictionary for a product name. Less than one-third of the unmatched barcodes (about 4.9% of total food items reported) were assigned an item description based on receipts and the rest of the unmatched barcodes (about 11.2% of total food items reported) did not have item descriptions. These food items presented a challenge for price imputation.

Westat believes that the performance of the scanner can be improved if respondents are given timely feedback on whether or not the scanned UPCs can be matched to the UPC data dictionary. Respondents can be further instructed to either scan the barcodes in the Primary Respondent's Food Book or enter an item description for unmatched UPCs. This will reduce the amount of post-survey reviewing and editing required for FoodAPS-2. Westat also recommends that ERS use FoodAPS-1 data to augment and update the UPC and the IRI data dictionary so as to increase the matching for FoodAPS-2. Again, this type of timely feedback is feasible with the alternative food collection system and hard to implement for paper food books. As mentioned earlier, Westat recommends a system that can provide timely feedback to respondents on the matching status of scanned UPCs.

Furthermore, about 31.5 percent of total food items reported do not have a scanned barcode before matching to the IRI dictionary. It is not clear whether it is because respondents did not scan the barcode or the barcodes are not available. To further improve the performance of the scanner, ERS needs to encourage respondents to scan when possible and to add more pictures of food items to the Primary Respondent Food Book. ERS is concerned about the increased burden arising from adding more pictures to the Primary Respondent Food Book. It could be a cumbersome task for respondents to flip through pages of items to find the right product item. Westat agrees that there is a trade-off between burden and the loss of information. However, with paper food books, there is

not much else one can do in this situation. An alternative system could be explored that can streamline the task for finding the barcode associated with the right product through autofill and type edit features. For instance, one can type "Ground Beef" to the alternative system; either the barcode associated with product ground beef is brought up to be scanned or the product description is recorded into the system directly.

5.4.2 How Many UPCs Could Be Identified?

As mentioned earlier, scanned barcodes are first matched to the UPC data dictionary, which combines information from Nielsen, Gladson, and Gregg London dictionaries into a single unduplicated list of barcodes consisting primarily of Nielsen data. According to the End of Survey Memo, about 193,693 food items were reported, yielding 88,583 unique food items. Among the unique food items, 33 percent have a scanned barcode matched to the UPC data dictionary and 21 percent have a scanned barcode that cannot be matched to the UPC data dictionary before the IRI was made available. The rest of the food items (46%) are not associated with a barcode, either because there is no barcode or respondents did not scan the barcode.

ERS later provided the IRI data dictionary. The preliminary match conducted by Mathematica found that 65 percent of scanned barcodes that did not match the Nielsen, Gladson, and London data but were assigned an item description from the receipts were found in the IRI data dictionary. Forty percent of scanned barcodes that were not matched with an item description can be found in the IRI database. Furthermore, 80 percent of scanned barcodes that were matched to the UPC data dictionary could be linked to IRI, which provides better standardization than the combination of Nielsen, Gladson, and London data. Of course, this linkage is important only if the resultant item descriptions are either better for getting nutrition characteristics or better for imputing price. If the objective is simply to capture price and description, this additional matching to IRI when the barcodes are already matched to the UPC data dictionary is not necessary.

In addition, some barcodes were dropped because they could not be matched to the UPC data dictionary or receipts were retrieved to be matched to the IRI.

Review of the Telephone Instrument 6

On Days 2, 5, and 7 of the study week, primary respondents were instructed to collect all of the household's food books and place a food reporting call to the telephone center at Mathematica. Respondents are given $10 for each of the three calls they initiated. If an inbound call was not received on the scheduled day, interviewers placed outbound calls beginning at noon the following day. During these food reporting calls, telephone interviewers asked the primary respondents to report all acquisitions written on the Daily Lists, including FAH and FAFH acquisitions. Then the telephone interviewer asked the primary respondents to report from all household members' Red Pages.

6.1 Combining or Replacing Telephone Interviews with an Alternative Food Reporting System

Telephone interviewers entered information from the Daily Lists into a real-time data entry system, including day of the acquisition, name of the place where food was acquired, who acquired the food, the total paid, and whether the acquisition was free. Detailed information from the Red Pages and/or attached receipts was also entered into the same data entry form in real time.

At the end of the study week, Red Pages were collected by field interviewers (together with Blue Pages) and sent back to Mathematica for processing. All Red Pages went through a "Red Page Review and Capture" process where information from physical Red Pages was first compared to what was entered by telephone interviewers. Pages whose information was not found in the system entered by telephone interviewers were data entered into the system. No doubt the review-and-capture process is labor intensive and costly.

Westat recommended earlier that ERS consider an alternative food reporting system that can be combined with the telephone interviewing (or possibly eliminate the need for telephone interviews) to substantially reduce the reporting burden on respondents' part. Such a system may reduce or eliminate the need for the review-and-capture process, dramatically reducing the post-survey processing burden on the survey organization's part. Instead, telephone interviewers could be charged to contact respondents to remind them of food acquisition entries when the alternative

food reporting system shows that respondents have not logged food acquisitions for any day or have missing data to certain data fields. Interviewers could remind respondents by calling them directly, sending text messages to their cell phones (upon their consent), or sending email messages (again upon their consent).

6.2 Examining the Telephone Instrument Design Features

Westat reviewed the Telephone Interview instrument. Step 1 of the instrument establishes the purpose of the call whereas Step 2 confirms and updates, if necessary, household size and household roster. Step 3 collects information from the Daily Lists. Questions Q4 to Q6c aim to capture FAFH acquisitions whereas Q7 to Q8f intend to capture FAH acquisitions. Q9 verifies whether or not the household member truly did not purchase any food item when his/her book does not contain any entry to the Daily Lists pages. Step 4 collects detailed information recorded on Red Pages. Telephone interviewers remind respondents to save receipts and fill out the Meals and Snacks Form at Step 5.

If ERS decides to continue with the Telephone Interviews as in FoodAPS-1, Westat suggests that ERS program in logic edits to check for potential redundancy (e.g., a meal shared by all household members at a restaurant is reported on multiple food books); inconsistency (e.g., the same meal was reported with different dollar amounts at the Daily Lists section and the Red Pages); and out-of-range values (e.g., a meal purchased at the McDonald's costs more than $100) in reports of food acquisitions.

6.3 Analyzing Telephone Interview Data to Examine Information Missed

Westat examined the telephone call log data (nfs_calllog.sas7bdat) to further the understanding of the amount of information missed by telephone interviews. We limited our analyses in this section (and the next two sections) to the 4,826 households that completed both the Initial and the Final Interview. Twenty households are not found in the telephone call log paradata and are removed from the analyses, yielding a total number of 4,806 households for the analyses. These households completed both CAPI interviews and were found in the telephone call log data.

Among the 4,806 households, 12.8 percent of them did not complete the Day 2 reporting (they did not call in and were also not reached through outbound calls). Approximately one-fifth (21.9%) of them did not complete the Day 5 reporting and 15.8 percent did not complete the Day 7 reporting. Looking at households' participation status in all 3 days' reporting, 63.6 percent of households (i.e., 3,059 households) completed the Telephone Interviews for all 3 days and 24.1 percent competed 2 days' interviews. One in 10 (10.3%) participated only in one day's telephone interviews and about 2 percent did not participate in any day's telephone interviews.

Note that if a household missed the Day 2 interview, the next call (e.g., the Day 5 interview) was supposed to pick up food events missed on Day 2 and to capture new food events since Day 2. If the Day 7 call was missed, the field interviewers were supposed to instruct the primary respondents to call the telephone center to catch up on reporting events before they could start the Final Interview. However, the results reported in this section and the next two sections are about households' participation status in telephone interviews rather than the coverage of food events captured by telephone interviews.

We looked at household participation status by key household characteristics that are both related to response propensity and household expenditures. Specifically, we looked at household size, income, and urbanicity. We also examined households' participation status in the three Telephone Interviews by these household characteristics. This analysis is weighted by household weights (revised household weights provided by Westat) and has taken into consideration complex survey design features (clustering and stratification).

For household size, we divided all 4,086 households into three groups based on the "hhsize" variable (located in a household level dataset named "household_imp.sas7bdat"): 1,019 households are single-person households, 2,201 households have two or three persons, and 1,586 households have four or more people. We found that about six in ten households with four or more people (64.4%) participated in all three Telephone Interviews, significantly less than single-person households (69.8%, Rao-Scott $\chi^2(1)=3.7$, $p=0.05$) and households with two to three people (73.0%, Rao-Scott $\chi^2(1)=9.4$, $p=0.002$).

We divided households into three groups based on the variable "inchhpovguide" (from the same household file): 1,521 households with income below 100 percent of the Federal poverty guideline, 1,256 households with income between 100 percent and 185 percent of the Federal poverty guideline, and 2,029 households with income greater than 185 percent of the Federal poverty guideline. Close to two-thirds of households with income below 100 percent of the Federal poverty

guideline (64.4%) participated in all three Telephone Interviews. A comparable proportion of households with income between 100 percent and 185 percent of the Federal poverty guideline (68.6%) did so; the difference between these two groups is marginally significant (Rao-Scott $\chi^2(1)=3.1$, $p=0.08$). By contrast, 72.1 percent of households with income greater than 185 percent of the Federal poverty guideline participated in all three Telephone Interviews, a proportion significantly higher than that for the households with income less than 100 percent of the Federal poverty guideline (Rao-Scott $\chi^2(1)=9.5$, $p=0.002$).

For the urbanicity indicator, we used the variable "uatype" from the same household file. The variable has three categories: urban (2,660 households), suburban (231 households), and rural (1,915 households). A total of 73.6 percent of rural households completed all three Telephone Interviews, significantly more than urban households (67.1%, Rao-Scott $\chi^2(1)=11.4$, $p=0.001$). About two-thirds of suburban households (66.8%) participated in all three telephone calls.

The analysis indicates that large households with four or more members are less likely to complete all three calls whereas households with income above 185 percent of the Federal poverty guideline and rural households are more likely to do so. In light of these findings, ERS should consider strategies to increase the participation rate for large households, households with income less than 100 percent of the Federal poverty guideline, and urban households if telephone calls are considered for FoodAPS-2. Possible strategies include increased incentive for households with lower participation rates and more upfront reminders.

As mentioned earlier, 20 cases were excluded from our analyses because they were not found in the telephone log data. Among the 20 cases, 10 of them have 2 or 3 persons in the household, five are single-person households and the remaining five are in large households with 4 or more people. In terms of household income, 4 of the 20 households have income below 100 percent of the Federal poverty guideline. Eight have income between 100 and 185 percent of the Federal poverty guideline and eight have income above 185 percent of the Federal poverty guideline. Seventeen cases are from urban and three from rural areas.

6.4 Characteristics of Respondents Who Initiated Telephone Calls

Among the 4,712 households that completed at least one of the three Telephone Interviews, 27.5 percent of them initiated the call for all three interviews. A little over one-third of them initiated the call for at least two of the Telephone Interviews and 26.4 percent initiated the call for only one of the Telephone Interviews. A total of 11.0 percent did not initiate the call for any interview.

We next examined the characteristics of respondents who initiated all three telephone calls. We found that large households with four or more people are significantly less likely to initiate all three calls (25.7%) than single-person households (32.1%, Rao-Scott $\chi^2(1)=7.2$, $p=0.007$). Close to one-third of households with two or three people (29.7%) initiated all three calls, a proportion that is marginally higher than that for large households with four or more people (Rao-Scott $\chi^2(1)=2.8$, $p=0.09$). Furthermore, there is not much difference in the proportion of households initiating all three telephone calls by income and urbanicity. The results suggest that, if ERS considers implementing the Telephone Interviews for FoodAPS-2, effort needs to be spent on encouraging large households to initiate the telephone calls.

Review of the Respondent Feedback Form 7

The Respondent Feedback Form is a paper-pencil form that is given to the respondent during the Final Visit. Respondents are asked to complete the form and enclose it in a sealed envelope. A total of 97.7 percent of respondents completed the feedback form. The form is key-entered on a flow basis throughout the field period. Item nonresponse ranged from 1 percent to 2 percent.

7.1 Possibility to Combine the Form with the Final Interview

Westat suggests that ERS administer the Respondent Feedback Form as an Audio Computer-Assisted Self-Interviewing (ACASI) instrument. The way an ACASI instrument works is like this: At the end of the CAPI interview and before the start of the ACASI instrument, interviewers will turn the laptop computer (or tablet computer) around to face the respondents. Interviewers will explain that respondents are supposed to answer some questions by entering their answers directly to the computer by themselves. They can see the questions on the computer screen or listen to the audio-recordings of the survey questions. Interviewers will be around and help only when asked. ACASI instruments have been shown to increase honest responding (Tourangeau and Smith, 1996; Tourangeau and Yan, 2007). Survey literature demonstrates that, in general, self-administered modes of data collection (including paper-pencil surveys, ACASI surveys, web surveys) are more effective than interviewer-administered modes of data collection (e.g., CAPI surveys) in reducing socially desirable responses and obtaining truthful answers. However, there is no evidence indicating significant differences among various forms of self-administered modes (Tourangeau and Yan, 2007). The three meta-analyses reported in Tourangeau and Yan (2007) fail to find an advantage associated with paper-pencil forms over computerized self-administration (such as ACASI or web surveys). The advantages of an ACASI instrument, compared to a paper-pencil instrument, are the faster turnaround of data and savings resulting from the elimination of the data entry process.

7.2 Possibility to Make the Form a Web Instrument

Making the Respondent Feedback Form a web instrument has the same advantages – faster turn-around of data and savings resulting from the elimination of data entry process – as the ACASI

instrument. However, the timing of fielding this web instrument is challenging. Asking respondents to complete the web feedback form *before* the Final Interview may change respondents' food acquisition behavior, whereas asking respondents to complete this web feedback form *after* the Final Interview may hurt response rates to the feedback form. Interviewers could ask respondents to log in and complete the web interview during the Final Visit. But Internet connectivity and computer or smartphone ownership might become a hindrance. Westat believes that an ACASI feedback instrument is a better option than a web feedback instrument.

7.3 Analysis of Q4

Q4 asks how respondents changed the way they got food because of their participation in FoodAPS-1. The question is a check-all-that-apply item offering six response options, an "Other changes – please specify" option, and a "No, did no change" option. Westat reviewed, in Section 7.3.1, open-ended responses for all 119 feedback forms that checked the "Other changes – please specify" box and wrote verbatim answers. In Section 7.3.2, Westat reviewed response distributions of the first response options.

7.3.1 Open-Ended Responses

"Among the 119 feedback forms, Mathematica recoded 19 of them back to the response list. The rest of the 100 feedbacks are considered "not codeable." Westat went through the 100 forms and found that 70 of them mentioned reasons for the study week to be different from a typical week (e.g., holidays, being sick, visitors, out of town/on vacation, going back to school/to work, more work, etc.). These households did eat out more or eat less, but the change is not due to participation in the study but to other events or changes in schedule. Thirteen people mentioned changes in other behaviors; for instance, they remembered to keep receipts or paid more attention to what they bought and what they ate. Six people mentioned the study design (e.g., did not like to call in). The rest of open-ended responses can't be easily coded. Based on this review, Westat does not think new options should be added to Q4.

7.3.2 Affirmative Responses to Q4_1 to Q4_4

Westat examined the respondent feedback data and found that 64 cases checked both the "Did not change" box (FF4_8) and at least one of the six boxes before the "Other changes – please specify" box. One respondent checked all boxes to Q4. When exploring a reasonable way to quantify the degree of changes, Westat excluded from analysis this case that checked all boxes to Q4 and reset the value of FF4_8 to 0 for the other 63 cases.

Westat also found other cases that provide apparently implausible answers to Q4_1 to Q4_4. For instance, one case answered positively to all four questions. Five cases checked both "Did More Shopping" and "Did Less Shopping" whereas four cases checked both "Ate Out More Often" and "Ate Out Less Often." Westat excluded these 10 cases from the analysis.

Among the rest of the 4,653 cases, Westat then counted the number of times respondents provided affirmative answers to Q4_1 to Q4_4. About 9 in 10 respondents (4,204 cases) did not answer affirmatively to any of these four questions. Among the remaining 449 cases, 100 answered positively only once to Q4_4 ("Did Less Food Shopping") and 107 only once to Q4_2 ("Ate Out Less Often"). Thirty-three cases provided two affirmative answers by indicating that they ate out less often and they did less food shopping. By contrast, 83 cases answered positively only once to Q4_3 ("Did More Food Shopping"), 74 cases only once to Q4_1 ("Ate Out More Often"), and 7 cases indicated that they ate out more often and they did more food shopping. Forty-five cases also provided two positive answers – 21 of them indicated that they ate out less often but did more food shopping, whereas 24 indicated that they ate out more often but did less food shopping.

Overall, given that only 7.8 percent of the cases (364) provided one affirmative answer to these four questions and 1.8 percent (85) cases provided two affirmative answers, Westat doesn't think there is enough variation to construct an index to quantify the degree of changes due to participation in FoodAPS-1.

Assessment of Response Burden and Response Rates 8

8.1 Response Rates to Telephone Interviews and Final Interview by Household Characteristics

For this analysis, Westat made use of existing variables in the screener dataset (nfs_screeners.sas7bdat). The purpose is *not* to replicate response rates published in FoodAPS-1 documentations. Instead, the purpose is to study the impact of response burden on respondents' propensity to participate in different tasks. We limited this analysis to households that were screened in as eligible and agreed to participate in the Initial Interview (status=020). Assuming all households are eligible to the later survey requests (the Initial Interview, Telephone Interviews, the Final Interview), response rates to these survey requests are simply the number of these people who completed the survey requests divided by those who were screened in as eligible and agreed to participate in the Initial Interview (n=6,185). Response burden is measured through the household size variable ("rhhsize" in the screeners dataset), assuming more burdensome reporting tasks for larger households, and target group ("rquotagroup" in the screeners dataset), assuming more burdensome reporting tasks for households below 100 percent of the Federal poverty guidelines. Households with missing data to "rhhsize" variable (n=5) and "rquotagroup" variable (n=43) were removed from the analysis.

Table 8-1 presents weighted response rates to the Initial Interview, weighted response rates to the Final Interview conditional on completing the Initial Interview (i.e., weighted proportion of Initial Interview completes that also completed the Final Interview), and weighted proportion completing all three telephone calls conditional on completing both the Initial and Final Interviews. The weights used in this set of analysis are household base weights that adjusted for unknown occupancy status and screener completion ("scr_comp_adj2_wgt" in the weights dataset prepared by Westat). Eleven cases were removed from analyses below because of non-positive weights. Pair-wise comparisons are conducted to test response rates across types of households.

The trend is apparent in Table 8-1; large households with at least four people have significantly a lower response rate to the Initial Interview and the Final Interview and a significantly lower rate of participation in all three telephone calls. As recommended earlier, ERS needs to think about ways to

encourage large households to participate in various tasks such as assigning best interviewers, allowing additional contact and recruitment effort, and providing more incentives.

Table 8-1. Weighted response rates to later survey requests, by household size

	Sample size	Weighted response rate to initial interview	Weighted response rate to final interview conditional on completing initial interview	Weighted proportion completing all three telephone calls conditional on completing both initial and final interviews
Total	**6,169**	**4,891**	**4,710**	**2,989**
A. Single-person households	1,246	79.6%	96.3%	69.5%
B. Households with 2 to 3 people	2,864	79.2%	97.2%	69.8%
C. Households with 4 or more people	2,059	75.2%	95.2%	61.0%
Significance tests				
A vs. B		$\chi(1)$=0.03, p=0.87	$\chi(1)$=0.42, p=0.52	$\chi(1)$=0.02, p=0.90
A vs. C		$\chi(1)$=3.64, p=0.06	$\chi(1)$=0.42, p=0.51	**$\chi(1)$=12.19**, p=0.001
B vs C		**$\chi(1)$=6.38**, p=0.012	**$\chi(1)$=7.88**, p=0.005	**$\chi(1)$=20.93**, p<0.001

Note: Weighted percentages were computed using scr_comp_adj2_wgt from dataset (wgt_factors_new.sas7bdat) and standards errors were computed using Taylor Series linearization. Rao-Scott Chi-Square values are reported in the table. Bolded Chi-Squares are statistically significant after the Bonferroni correction.

Table 8-2 displays the same (weighted) rates and (weighted) proportions by target group. Again, pair-wise comparisons are conducted to test response rates across types of households.

Table 8-2 suggests that households in Group D (Households receiving SNAP benefits) have a significantly higher response rate to the Initial Interview than those in Group B (Households not receiving SNAP benefits and whose income is between 100 and 185 percent of the Federal poverty guideline). Conditional on completing the Initial Interview, Group C (Households not receiving SNAP benefits and whose income is greater than 185 percent of the Federal poverty guideline) has a significantly higher response rate to the Final Interviews than Group A (Households not receiving SNAP benefits and whose income is below 100 percent of the Federal poverty guideline). Furthermore, Group C has a significantly higher proportion participating in all three telephone interviews than the other three types of households. This finding suggests that ERS may want to consider strategies to increase participation of low-income households in telephone interviews if

telephone interviews are to be kept for FoodAPS-2. Potential strategies include assigning best interviewers, providing additional incentives to participate, and adding more timely reminders.

Table 8-2. Weighted response rates to later survey requests, by target group

	Sample size	Weighted response rate to initial interview	Weighted response rate to final interview conditional on completing initial interview	Proportion completing all three telephone calls conditional on completing both initial and final interviews
Total	**6,131**	**4,852**	**4,671**	**2,965**
A. Households not receiving SNAP benefits and whose income is below 100 percent of the Federal poverty guideline	1,047	76.9%	93.9%	57.4%
B. Households not receiving SNAP benefits and whose income is between 100 and 185 percent of the Federal poverty guideline	2,023	76.7%	96.6%	66.1%
C. Households not receiving SNAP benefits and whose income is greater than 185 percent of the Federal poverty guideline	1478	77.7%	97.8%	75.6%
D. Households receiving SNAP benefits	1583	82.3%	95.8%	62.8%
Significance tests				
A vs. B		$\chi(1)$=0.01, p=0.91	$\chi(1)$=3.01, p=0.08	$\chi(1)$=11.10, p=0.009
A vs. C		$\chi(1)$=0.12, p=0.73	**$\chi(1)$=8.51**, p=0.004	**$\chi(1)$=48.16**, p<0.0001
A vs. D		$\chi(1)$=5.54, p=0.02	$\chi(1)$=1.25, p=0.26	$\chi(1)$=3.66, p=0.06
B vs. C		$\chi(1)$=0.28, p=0.60	$\chi(1)$=3.42, p=0.06	**$\chi(1)$=20.49**, p<0.0001
B vs. D		**$\chi(1)$=7.25**, p=0.007	$\chi(1)$=0.88, p=0.35	$\chi(1)$=1.99, p=0.16
C vs. D		$\chi(1)$=5.03, p=0.02	$\chi(1)$=6.45, p=0.01	**$\chi(1)$=29.32**, p<0.0001

Note: Weighted percentages were computed using scr_comp_adj2_wgt from dataset (wgt_factors_new.sas7bdat) and standards errors were computed using Taylor Series linearization. Rao-Scott Chi-Square values are reported in the table. Bolded Chi-Squares are statistically significant after the Bonferroni correction.

Both Table 8-1 and Table 8-2 provide evidence for the association between response burden and likelihood of respondents participating in survey requests. ERS is interested in knowing the relationship between burden and response rates to the Initial Interview, Final Interview, and the three telephone interviews when considering both household size and income level. Logistic

regression models are run to predict individual respondents' likelihood to participate in the Initial Interview, Final Interview, and all three telephone calls as a function of household size, target group, and the interaction of those two. The results of the model predicting propensity to participate in the Initial Interview indicate that both household size and target group have a statistically significant main effect on respondents' likelihood to participate in the Initial Interview, but the interaction effect is not statistically significant. The results of the model predicting propensity to participate in the Final Interview show a marginally significant main effect of target group. The results of the model predicting the likelihood to participate in all three telephone interviews reveal a significant main effect of household size and a marginally significant interaction of household size and target group.

In addition, ERS is interested in the burden of households that are both low income and large. For this purpose, we recoded the household size variable to contrast large households (with four or more people) to small households (with less than four people) and recoded the income variable to contrast low-income households (including households receiving SNAP benefits and households not receiving SNAP benefits and whose income is less than 185 percent of the Federal poverty guideline) to high-income households (with income greater than 185 percent of the Federal poverty guideline). In Table 8-3, we present weighted response rates for each of the four groups after crossing household size and income. We contrast the other three groups to households that are both low income and large.

It is apparent from Table 8-3 that households that are both low income and large have lower response rates to the Initial Interview, but significantly lower conditional response rates to the Final Interview than small households with a high income. Large households with a low income have the lowest proportion participating in all three telephone interviews. Table 8-3 suggests that households that are both low income and large have lower propensity to participate in later survey requests; thus, ERS should consider spending more effort on this type of household to ensure adequate representation.

Table 8-3. Weighted response rates to later survey requests, by household size and income

	Sample size	Weighted response rate to initial interview	Weighted response rate to final interview conditional on completing initial interview	Proportion completing all three telephone calls conditional on completing both initial and final interviews
Total	**6,129**	**4,851**	**4,670**	**2,964**
A. Low-income and Large	1,701	75.3%	94.5%	58.1%
B. High-income and Small	1,131	78.6%	97.9%	77.2%
C. Low-income and Small	2,951	79.5%	96.3%	65.6%
D. High-income and Large	346	74.5%	97.5%	69.5%
Significance tests				
A vs. B		$\chi(1)$=2.42, p=0.12	**$\chi(1)$=12.65**, p=0.0004	**$\chi(1)$=59.44**, p<0.0001
A vs. C		$\chi(1)$=5.49, p=0.02	$\chi(1)$=2.92, p=0.09	**$\chi(1)$=12.67**, p=0.0004
A vs. D		$\chi(1)$=0.08, p=0.78	$\chi(1)$=4.34, p=0.04	**$\chi(1)$=8.76**, p=0.003

Note: Weighted percentages were computed using scr_comp_adj2_wgt from dataset (wgt_factors_new.sas7bdat) and standards errors were computed using Taylor Series linearization. Rao-Scott Chi-Square values are reported in the table. Bolded Chi-Squares are statistically significant after the Bonferroni correction.

8.2 Diary Entries by Household Characteristics

Westat focused on the total number of food events and the total number of food items reported by households as well as the number of food events and food items per person (the total number divided by the household size) for this analysis. We again restricted analyses to the 4,826 households that completed both the Initial Interview and the Final Interview. Food data come from the two item level food datasets (faps_fahitem.sas7bdat and faps_fafhitem.sas7bdat) and household characteristics are taken from the same household level data as mentioned in Section 6.3.

For each household, Westat counted the number of food events reported and the number of food items reported to derive the total number of food events and food items reported. We then divided the total numbers by household size to produce per-person numbers as standardized measures. Table 8-4 shows the weighted means of both total number at the household level and per-person numbers by food type and household size. Weights used in this analysis are final household weights (variable "hhwgt" in the faps_hhweightsnew.sas7bdat) prepared by Westat. Again, pair-wise comparisons are conducted to examine food events and food items across types of households.

Table 8-4. Weighted mean number of food events and food items reported at the household level and person level, by food type and household size

	Weighted mean number of FAH events reported		Weighted mean number of FAH items reported		Weighted mean number of FAFH events reported		Weighted mean number of FAFH items reported	
	Per household	Per person	Per household	Per person	Per household	Per person	Per household	Per person
Overall mean	3.5	1.6	32.6	13.7	8.5	3.4	26.7	10.4
A. Single-person households	2.7	2.7	20.3	20.3	4.7	4.7	14.0	14.0
B. Households with 2 to 3 people	3.5	1.5	31.4	13.6	7.6	3.2	23.9	10.1
C Households with 4 or more people	4.0	0.8	43.5	9.0	12.6	2.6	39.9	8.4
Significance tests								
A vs. B	**t=-6.84**, p<0.0001	**t=15.28**, p<0.0001	**t=-11.09**, p<0.0001	**t=8.61**, p<0.0001	**t=-10.23**, p<0.0001	**t=6.67**, p<0.0001	**t=-9.32**, p<0.0001	**t=4.23**, p<0.0001
B vs. C	**t=-4.05**, p=0.0002	**t=14.08**, p<0.0001	**t=-8.26**, p<0.0001	**t=12.00**, p<0.0001	**t=-12.38**, p<0.0001	**t=5.85**, p<0.0001	**t=-11.13**, p<0.0001	**t=5.01**, p<0.0001
C vs. A	**t=11.79** p<0.0001	**t=-28.49**, p<0.0001	**t=15.57**, p<0.0001	**t=-14.78**, p<0.0001	**t=18.83**, p<0.0001	**t=-9.43**, p<0.0001	**t=16.33**, p<0.0001	**t=-6.12**, p<0.0001

Note: Weighted percentages were computed using hhwgt and standard errors were computed using Taylor Series linearization. Bolded T-values are statistically significant after the Bonferroni correction.

We found that, with regards to the total number of food events and food items reported at the household level, there is a linear trend. Large households with four or more people reported significantly more events and more items than households with two to three people, who reported significantly more events and more items than single-person households for both food at home and food away from home, confirming that large households do have a more burdensome reporting task. However, large households reported significantly fewer per-person food events and food items than smaller households. For instance, single-person households reported an average of 2.7 FAH events per person whereas large households with four or more people reported only an average of 0.8 FAH events per person. This might suggest that large households underreported food events and food items at the person level despite reporting higher overall numbers. Westat suggests that ERS look into the variations at the per-person level to determine whether the trend in Table 8-4 is reasonable.

Displayed in Table 8-5 are weighted mean numbers of food events and items at both the household level and per-person level reported by income group. (Income groups are defined in the same manner as in Section 6.3. SNAP households are included in this analysis as well.) There is not much differences between levels of income in terms of FAH events and FAH items reported. However,

households with income greater than 185 percent of the Federal poverty guideline reported significantly more FAFH events and FAFH items at household and per-person levels than households with income less than 100 percent of the Federal poverty guideline, and significantly more FAFH events at household and per-person levels and significantly more per-person FAFH items than households with income between 100 and 185 percent of the Federal poverty guideline. Table 8-5 does not seem to indicate that low-income households have a higher reporting burden than high-income households in terms of number of food items and food events reported. Given that low-income households are less likely to complete all three telephone interviews, the fewer food items reported by low-income households might reflect underreporting by these households.

Table 8-5. Weighted mean number of food events and food items reported at the household level and per-person level, by food type and income

	FAH events reported (Weighted mean)		FAH items reported (Weighted mean)		FAFH events reported (Weighted mean)		FAFH items reported (Weighted mean)	
	House-hold	Per person	House-hold	Per person	House-hold	Per person	House-hold	Per person
A. Households with income below 100% of Federal poverty guideline	3.4	1.6	32.5	13.2	7.5	2.8	23.9	8.8
B. Households with income between 100% and 185% of Federal poverty guideline	3.5	1.6	30.3	13.2	8.2	3.1	25.5	9.5
C. Households with income above 185% of Federal poverty guideline	3.5	1.5	33.9	14.2	9.2	3.7	28.6	11.5
Significance tests								
A vs. B	*t*=-0.23, *p*=0.41	*t*=-0.88, *p*=0.19	*t*=1.44, *p*=0.08	*t*=0.05, *p*=0.48	*t*=-1.28, *p*=0.10	*t*=-2.19, *p*=0.02	*t*=-1.00, *p*=0.16	*t*=-1.48, *p*=0.07
B vs. C	*t*=0.12, *p*=0.45	*t*=1.29, *p*=0.10	***t*=-2.74,** *p*=0.005	*t*=-1.31, *p*=0.10	***t*=-2.57,** *p*=0.008	***t*=-3.94,** *p*=0.0002	*t*=-2.32, *p*=0.01	***t*=-4.06,** *p*=0.0001
C vs. A	*t*=0.14 *p*=0.45	*t*=-0.27, *p*<0.39	*t*=1.24, *p*=0.11	*t*=-1.91, *p*=0.03	***t*=3.98,** *p*=0.0002	***t*=6.09,** *p*<0.0001	***t*=3.53,** *p*=0.0006	*t*=5.12, *p*<0.0001

Note: Weighted percentages were computed using hhwgt and standard errors were computed using Taylor Series linearization. Bolded T-values are statistically significant after the Bonferroni correction.

Reports of food acquisition by urbanicity are presented in Table 8-6. There is no significant difference in terms of food events and food items by location of households, even though rural households are more likely to participate in all three telephone interviews than urban households

(see Section 6.3). There doesn't seem to be much variation in reporting burden in terms of number of food events and items reported.

Table 8-6. Weighted mean number of food events and food items reported at household level and person level, by food type and urbanicity

	FAH events reported (Weighted Mean)		FAH items reported (Weighted Mean)		FAFH events reported (Weighted Mean)		FAFH items reported (Weighted Mean)	
	House-hold	Per person	House-hold	Per person	House-hold	Per person	House-hold	Per person
A. Urban households	3.6	1.5	32.7	13.2	8.7	3.3	26.3	10.0
B. Suburban households	3.6	1.4	32.6	11.8	8.9	3.1	27.3	9.5
C. Rural households	3.3	1.6	32.5	14.4	8.3	3.4	27.1	11.0
Significance tests								
A vs. B	*t*=0.07, *p*=0.47	*t*=1.00, *p*=0.16	*t*=0.04, *p*=0.48	*t*=1.03, *p*=0.16	*t*=-0.27, *p*=0.40	*t*=0.59, *p*=0.28	*t*=-0.42, *p*=0.34	*t*=0.40, *p*=0.35
B vs. C	*t*=1.33, *p*=0.10	*t*=-1.52, *p*=0.07	*t*=0.01, *p*=0.50	*t*=-2.01, *p*=0.03	*t*=0.66, *p*=0.26	*t*=-0.74, *p*=0.23	*t*=0.09, *p*=0.46	*t*=-1.13, *p*=0.13
C vs. A	*t*=-2.87 *p*=0.004	*t*=0.75, *p*=0.23	*t*=-0.22, *p*=0.42	*t*=2.32, *p*=0.01	*t*=-0.76, *p*=0.23	*t*=0.42, *p*=0.34	*t*=0.51, *p*=0.31	*t*=1.87, *p*=0.04

Note: Weighted percentages were computed using hhwgt and standard errors were computed using Taylor Series linearization. Bolded T-values are statistically significant after the Bonferroni correction.

8.3 Length of Interview by Household Characteristics

Westat also examined the length of various interviews as a second indicator of burden. The interview for the Initial Interview and the Final Interview are taken from the timings dataset (nfs_capitimes.sas7bdat) and the length of telephone interviews are taken from the telephone paradata (nfs_calllog.sas7bdat). Household characteristics are taken from the same household dataset as described in Section 6.3. Weights used in this analysis are final household weights (variable "hhwgt" in the faps_hhweightsnew.sas7bdat) prepared by Westat. Again, pair-wise comparisons are conducted to examine food interview length across types of households..

Table 8-7 displays the weighted mean interview length in minutes by interview and by household size. Not surprisingly, across all interviews, households with four or more people took a significantly longer time than households with two to three people (except for the Final Interview) and single-person households. Households with two or three people spent a significantly longer time than single-person households across all interviews. Table 8-7 confirms that large households have a

more burdensome reporting task than smaller households as indicated by the longer interview time and higher report of food events and food items at the household level (in Table 8-4).

Table 8-7. Mean interview length by interview type and household size

	Initial interview	Final interview	Day 2 telephone call	Day 5 telephone call	Day 7 telephone call
Overall mean	**15.1**	**24.7**	**10.8**	**10.9**	**8.3**
A. Single-person households	12.1	22.1	8.4	8.0	5.3
B. Households with 2 to 3 people	15.3	25.7	11.6	12.2	9.2
C. Households with 4 or more people	20.8	27.3	13.5	14.5	13.5
Significance tests					
A vs. B	***t*=-5.65,** *p*<0.0001	***t*=-3.21,** *p*=0.002	***t*=-6.07,** *p*<0.0001	***t*=-8.43,** *p*<0.0001	***t*=-8.25,** *p*<0.0001
B vs. C	***t*=-8.03,** *p*<0.0001	*t*=-1.61, *p*=0.06	***t*=-2.53,** *p*=0.008	***t*=-3.41,** *p*=0.001	***t*=-7.09,** *p*<0.0001
A vs. C	***t*=-13.26,** *p*<0.0001	***t*=-4.35,** *p*=0.0001	***t*=-6.85,** *p*<0.0001	***t*=-10.83,** *p*<0.0001	***t*=-12.52,** *p*<0.0001

Note: Weighted means were computed using hhwgt and standard errors were computed using Taylor Series linearization. Bolded T-values are statistically significant after the Bonferroni correction.

Table 8-8 presents interview length by interview type and by income. There is no significant difference across income levels in interview length for the two CAPI surveys (the Initial and the Final Interview). However, households with income greater than 185 percent of the Federal poverty guideline spent a significantly longer time than households with income below 100 percent of the Federal poverty guideline to complete the Day 2 call and the Day 5 call, and a significantly longer time than households with income between 100 percent and 185 percent of the Federal poverty guideline to complete the Day 5 call and the Day 7 call. In terms of interview length, households with more money have a more burdensome reporting task than households with less money, as they reported significantly FAFH events and items (shown in Table 8-5).

Table 8-8. Weighted mean interview length by interview type and income

	Initial interview	Final interview	Day 2 telephone call	Day 5 telephone call	Day 7 telephone call
Overall mean	**15.1**	**24.7**	**10.8**	**10.9**	**8.3**
A. Households with income less than 100 percent of Federal poverty guideline	15.7	23.2	9.7	9.0	7.7
B. Households with income between 100 and 185 percent of Federal poverty guideline	15.2	25.5	10.4	9.9	7.1
C. Households with income greater than 185 percent of Federal poverty guideline	14.8	25.0	11.2	11.9	8.7
Significance tests					
A vs. B	t=0.68, p=0.25	t=-1.68, p=0.05	t=-1.05, p=0.15	t=-1.47, p=0.08	t=1.28, p=0.10
B vs. C	t=0.47, p=0.32	t=0.52, p=0.30	t=-1.50, p=0.07	**t=-3.60, p=0.0005**	**t=-3.16, p=0.002**
A vs. C	t=1.46, p=0.08	t=-1.36, p=0.09	**t=-2.96, p=0.003**	**t=-4.71, p=0.0002**	t=-1.80, p=0.04

Note: Weighted means were computed using hhwgt and standards errors were computed using Taylor Series linearization. Bolded T-values are statistically significant after the Bonferroni correction.

As shown in Table 8-9, there is no significant difference across urbanicity in interview length, just as there is no significant difference across urbanicity in the number of food events and food items reported.

Table 8-9. Weighted mean interview length by interview type and urbanicity

	Initial interview	Final interview	Day 2 telephone call	Day 5 telephone call	Day 7 telephone call
Overall mean	**15.1**	**24.7**	**10.8**	**10.9**	**8.3**
A. Urban households	16.2	25.4	10.9	10.6	8.9
B. Suburban households	13.9	24.2	10.5	11.1	7.8
C. Rural households	14.1	24.0	10.7	11.2	7.7
Significance tests					
A vs. B	t=1.14, p=0.13	t=-0.21, p=0.42	t=0.53, p=0.30	t=-0.33, p=0.37	t=0.84, p=0.20
B vs. C	t=0.08, p=0.47	t=0.04, p=0.49	t=-0.20, p=0.42	t=-0.09, p=0.47	t=0.13, p=0.44
A vs. C	**t=2.96, p=0.003**	t=1.21, p=0.12	t=0.67, p=0.25	t=-1.39, p=0.09	t=2.22, p=0.02

Note: Weighted means were computed using hhwgt and standard errors were computed using Taylor Series linearization. Bolded T-values are statistically significant after the Bonferroni correction.

Assessment of Use of Incentives 9

In FoodAPS-1, a $5 unconditional incentive is offered to all households contacted during the screening interviews. For households that participated in the study and completed data collection, the following incentives are provided at the end of the data collection week:

1. The primary respondent receives a $100 check (the base incentive);
2. The primary respondent receives a $10 gift card for initiating each of the three scheduled telephone calls during the study week;
3. Children ages 11-14 years receive a $10 gift card if they filled out their Youth Food Books or otherwise reported food acquisitions to the primary respondent; and
4. Members 15 and older receive a $20 gift card if they filled out their Adult Food Book or otherwise reported food acquisitions to the primary respondent.

Based on this incentive structure, a single-person household receives up to $135 in incentives and a two-person household receives up to $155 in incentives. This incentive structure is designed to encourage both initial agreement to participate in the study and continued participation throughout the study week.

As described in the Data Quality Memo #1, the Initial Interview was completed by 5,023 households, with 4,826 completing the data collection week and the Final Interview. That translates to a response rate of 96.1 percent to the Final Interview conditional on completing the Initial Interview. Conditional on completing both the Initial and Final Interviews, the response rate to the Respondent Feedback Form is 97.7 percent and to the Meals and Snacks Form is 91.7 percent. In addition, 90.6 percent of households that completed both the Initial and Final Interviews filled out the Red Pages and 92.1 percent filled out Blue Pages, based on numbers reported in the End of Survey Memo. As discussed in Section 8.1, more than 97 percent of households that completed both the Initial and Final Interviews answered at least one telephone call. Overall, Westat believes that the current incentive scheme is effective at encouraging continued participation throughout the study week, and Westat suggests that the current incentive scheme be kept if ERS were to keep the same study design for FoodAPS-2.

Conclusions and Recommendations 10

Westat conducted an expert review on all instruments and forms used in FoodAPS-1 and described, in this report, our review of the instruments and forms, and presented our recommendations for each of the instruments and forms.

Looking ahead at FoodAPS-2, Westat had two major recommendations to reiterate here. First, Westat firmly believes that FoodAPS-2 should fully exploit the computerization of all instruments and forms and maximize the functionalities with computerizing all instruments and forms. For instance, the screener form can be computerized by making it a CAPI instrument (instead of a paper-pencil instrument). The Respondent Feedback Form can also be computerized by making it an ACASI instrument. The advantages of computerizing these instruments and forms are to have data readily available from all these pieces, to populate or pre-fill data across instruments, and to build in edit checks that reduce data anomalies and improve the quality of data obtained. Westat especially encourages ERS to work on the populating and filling of data across instruments as well as programmed edit checks. For instance, ERS can populate date, day of week, and household member names for the computerized Meals and Snacks Form. Westat has made suggestions throughout the report on edit checks to be programmed and utilized for FoodAPS-2.

Second, to replace paper food books and forms, Westat strongly recommends that ERS explore an alternative food reporting system that has the potential to eliminate the need for manual data entry, to reduce duplications across household members and/or food books, to reduce inconsistencies in reporting, and to provide timely feedback to respondents if needed to resolve issues around blank days or missed days.

Westat also conducted secondary data analyses to examine the impact of burden on response rates and diary entries. We found that large households do have more burdensome reporting tasks to begin with. They reported more food events and food items in the food books. They took a longer time to complete the Initial Interview, the Final Interview, and the telephone calls. They also have a lower likelihood to complete various tasks during the data collection week. For FoodAPS-2, Westat suggests that ERS be sensitive about the level of burden FoodAPS-2 has on large households and be proactive in engaging their continued participation throughout the study week.

References

Edgar, J. (2010). *Respondent record use in the U.S. Consumer Expenditure Interview Survey*. Paper presented at the Annual Conference of the American Association for Public Opinion Research, May 13-16, 2010, Chicago, IL.

Juster, T., and Smith, J.P. (1997). Improving the quality of economic data: Lessons from the HRS and AHEAD. *Journal of the American Statistical Association*, 92, 1268–1278.

Kopp, B. (2013). *Use of financial records in the CE survey*. Paper presented at the CE Methods Symposium.

Sjoblom, M., and Lee, L. (2012). *Pilot test report: Records Information and Feasibility of Use Study*. Chicago: NORC at the University of Chicago, July 9, 2012.

Tourangeau, R., Rips, L.J., and Rasinski, K. (2000). *The psychology of survey response*. Cambridge, UK: Cambridge University Press.

Tourangeau, R., Shapiro, G., Kearney, A., and Ernst, L. (1997). Who lives here? Survey undercoverage and household roster questions. *Journal of Official Statistics*, 13, 1-18.

Tourangeau, R., and Smith, T.W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly*, 60, 275-304.

Tourangeau, R., and Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133, 859-833.

West, B., Conrad, F., Kreuter, F., and Mittereder, F. (2016). Can conversational interviewing improve survey response quality without increasing interviewer variance? Manuscript submitted to *Journal of Statistical Society, Series A*.

Yan, T., Curtin, R., and Jans, M. (2010). Trends in income nonresponse over two decades. *Journal of Official Statistics*, 26, 145-164.